INHOMOGENEOUS LENS STRUCTURES FOR INTEGRATED OPTICS

A Thesis Submitted to

the Faculty of Engineering

of the University of Glasgow

for the degree of

Doctor of Philosophy

by

Neil Finlayson, B.Sc.

August, 1985

Do mo phàrantan

And God said, Let there be light:
and there was light.

THE FIRST BOOK OF MOSES, CALLED GENESIS
Chapter 1, verse 3

O what a bursting out there was,
And what a blossoming,
When we had all the summer-time
And she had all the spring!

"A MAN YOUNG AND OLD"

W.B. Yeats

ACKNOWLEDGEMENTS

I should like to thank Professor J. Lamb for the provision of research facilities for this project in the Department of Electronics and Electrical Engineering and for sterling supervision.

I am greatly indebted to Dr. R.M. De La Rue for his careful supervision, for his excellent and painstaking reading of this thesis and, not least, for his constant support. I thank him sincerely.

To Mr. S. Gratze and to Mr. A. Hodkin of the Marconi Research Centre, Chelmsford, England, I would like to express my thanks, especially for allowing me to report on work jointly carried out on the edge-rounded spherical lens.

The geodesic lens ray-tracing program reported in chapter four and the BPM programs reported in chapter five were initially created by Dr. J. van der Donk of the University of Gent, Belgium. I thank Dr. van der Donk and Professor P. Lagasse for allowing me to use and modify these programs.

Dr. G. Doughty, Dr. J. Singh and Mr. J. Smith were enthiusiastic and helpful colleagues and contributed considerably to much of the work reported herein. I wish to thank Dr. Doughty, in particular, for allowing me to use the photographs shown in Figures [8.22], [8.23] and [8.28] .

I gratefully acknowledge the help of Dr. G. Stewart in connection with the fabrication and control of As_2S_3 thin-films. Similarly, I should like to thank Dr. A. Dawar for his help with the formation of titanium-diffused waveguides.

Ms. A. Mackinnon and colleagues were patient and skilled advisors on the solution of many of my computing problems.

The assistance of Mr. P. Anderson and Ms. S.W. Wong, on

experimental and theoretical problems respectively, is gratefully acknowledged.

I wish to express my admiration and thanks to the entire technical staff of the Department, particularly to Mr. G. Boyle, Mr. T. Wright, Mr. H. Anderson, Ms. L. Hobbs, Mr. J. Clark, Mr. J. Young and Mr. K. Piechowiak. Mr. R. Hutchins is thanked for his advice, encouragement and inimitable good humour.

I have enjoyed the support of all my friends. Dr. D. Murray (my landlord) deserves special thanks for allowing me to mess up his home for so long.

The financial support of the Science and Engineering Research Council and GEC Marconi is gratefully acknowledged.

Finally, I thank my family, from Lathamor and Adabrock, in the district of Ness, in the Isle of Lewis.

ABSTRACT

The thesis is concerned with the design, analysis, fabrication and evaluation of integrated optic lenses which are inhomogeneous either in physical shape or in refractive index profile. The thesis has nine chapters. Chapter one, the introduction, illustrates the importance of these lenses within the domain of integrated optics, where the complicated mathematical functions required to describe the lens profiles are most easily realised. Connections are made between the study of these lenses and the exciting new field of optical computing.

A special class of non-uniform lenses which are conceptually perfect optical instruments forms the main area of interest in the present study. Historically, the development of these lenses has followed two distinct lines, related to two possible methods of physically obtaining the required variation in path of light rays passing through the lens. In one method the optical path is made to vary directly, whilst the other method involves controlling the physical path, and thus the optical path, through the principle of equivalence. The dual development has been continued in the field of integrated optics, where lenses based on direct control of the optical path are termed variable-index lenses and those based on physical path control are termed geodesic lenses. The perfect variable-index lens studied in this work was the well-known Luneburg lens. Perfect geodesic lens designs have also been published. The design formulae for both types of lens are presented in chapter two. A simpler lens, of spherical geometry, is also presented which is easily analyzed and characterised and which serves as an archetypal model against which the performance of the more sophisticated lenses can be assessed.

Chapter three investigates the problems involved in modelling fabrication conditions in a thermal-evaporation-in-vacuum environment so that lens profiles can actually be constructed. Chapter four goes into methods of tracing rays through these lenses in some detail. Ray-tracing has long been

the classical tool of optical designers, providing a useful guide to optical performance. Ray methods, which effectively provide image error evaluations, are not entirely appropriate for those lenses which are conceptually perfect within the geometrical optics approximation. Diffraction effects prevent the lenses from attaining true perfection. In such cases the wave-field produced by the lenses in the image space is the important quantity. In chapter five, the beam-propagation method (BPM) is used to study diffraction and associated effects in inhomogeneous lenses. The method allows the propagation of complicated waveshapes in inhomogeneous media, normally a difficult task. Furthermore, anisotropic effects and the interaction between acoustic waves and optical waves can be studied with the method. Negative focal shifts are reported which are not predicted by geometrical optics or the usual approximate diffraction theories.

The fabrication of lenses is considered in chapter six. Planar waveguide measurements carried out on the various materials used in the study are presented. A major problem in the fabrication of geodesic lenses, that of obtaining a uniform waveguide layer over the complete lens area, is dealt with in some detail in chapter seven. In chapter eight, extensive tests on the experimental performance of several lenses are reported. Near diffraction-limited performance is reported for geodesic lenses. More limited performance figures are obtained for Luneburg lenses though the possibility of high performance is indicated if profile resolution can be improved. The themes of the thesis are pulled together for discussion in chapter nine and conclusions are drawn as to the relative merits of the various lenses. Possible means of improving fabrication procedures, thus driving lenses closer to ultimate resolution limits, are presented. The greatest problem faced is that of scattering in the waveguide, which appears to be accentuated as the waveguide traverses the lens surface. If the scattering problem can be successfully dealt with it is concluded that integrated optical lenses could be important and viable components in addressing the problem of fast, high-throughput data processing.

TABLE OF CONTENTS

		Page
ACKNOWLEDGEMENTS		i
ABSTRACT		iii
CHAPTER 1 :	INTRODUCTION	1
	References	8
CHAPTER 2 :	THE DESIGN OF INHOMOGENEOUS LENSES	
2.1	Introduction	9
2.2	Examples	9
2.3	Two Lines of Development	11
	2.3.1 The principle of equivalence	11
	2.3.2 Variable-refractive-index lenses	14
	2.3.3 Configurational/geodesic lenses	18
2.4	Design of the Thin-film Luneburg Lens	19
	2.4.1 The generalized solution given by Luneburg	19
	2.4.2 Southwell's method	24
	2.4.3 An analytical approximation	25
	2.4.4 Refractive index profiles	27
	2.4.5 The thin-film overlay Luneburg lens	27
2.5	Design of the Aspheric Geodesic Lens	32
2.6	A Simple Geodesic Design	35
2.7	Conclusions	36
	References	37
CHAPTER 3 :	THE SHADOW-MASKING PROBLEM	
3.1	Introduction	43
3.2	Holland and Steckelmacher Models	44
3.3	Lens Fabrication Using a Directed-Surface Source	47
3.4	The Yao Shadow-Masking Model	50
3.5	Masking the Substrate	51

3.6 Lens Approximations 55
3.7 Conclusions 57
References 58

CHAPTER 4 : TRACING RAYS THROUGH INHOMOGENEOUS LENSES
4.1 Introduction 59
4.2 Pre-Tracing 60
4.3 Fundamental Equations of Geometrical Optics 61
4.4 Tracing Rays Through Overlay Lenses 66
4.4.1 Introduction 66
4.4.2 The Southwell-Montagnino method 67
4.4.3 Exact point of crossing 73
4.4.4 Interpolation procedure 74
4.5 Tracing Rays Through Geodesic Lenses 74
4.5.1 Southwell's method 74
4.6 Results 79
4.6.1 Introduction 79
4.6.2 Overlay lenses 79
4.6.3 Discussion 82
4.6.4 Geodesic lenses 83
4.7 Conclusions 86
References 87

CHAPTER 5 : THE BEAM-PROPAGATION METHOD
5.1 Introduction 89
5.2 The Angular Spectrum of Plane Waves 90
5.3 Computational Representation of One-Dimensional Wave-Field 92
5.4 Algorithm for Propagation in Homogeneous Space 94
5.5 Theory of the Beam-Propagation Method 95
5.6 Conditions for Applicability 104
5.7 The Sampling Interval 106
5.8 Implementation 107
5.9 Results 108
5.9.1 Group (a) 112

5.9.2 Group (b) 113
5.9.3 Group (c) 114
5.9.4 Group (d) 116
5.9.5 Group (e) 116
5.10 Discussion 117
5.11 Conclusions 120
References 121

CHAPTER 6 : FABRICATION OF WAVEGUIDES AND OVERLAY LENSES
6.1 Introduction 124
6.2 Materials 124
6.2.1 The substrate material 124
6.2.2 The ambient waveguide 127
6.2.3 Overlay lens material 128
6.3 Fabrication of planar waveguides 132
6.3.1 Ti:$LiNbO_3$ waveguides 132
6.3.2 As_2S_3 waveguides 136
6.3.2.1 Fabrication 136
6.3.2.1 Evaluation 136
6.3.2.3 Optical evaluation 137
6.4 Fabrication of Overlay Lenses 139
6.4.1 Deposition of lenses 139
6.4.2 Measurement of profiles 140
6.5 Conclusions 143
References 145

CHAPTER 7 : GEODESIC LENS WAVEGUIDES
7.1 Introduction 152
7.2 Factors Affecting Lens Transmission 152
7.3 Deposited Film Profiles 154
7.4 Conclusions 161
References 162

CHAPTER 8 : OPTICAL EXPERIMENTS
8.1 Experimental objectives 163
8.2 Lenses Selected for Study 164
8.3 Special Tests and Techniques 167

8.3.1 Foucault test 168
8.3.2 Hartmann test 169
8.3.3 Resolution test 172
8.3.4 Direct observation 174
8.4 Comparisons Between Experiment and Theory 175
8.5 Apparatus and Experimental Procedure 178
8.6 A note on results and discussion 181
8.7 Results - SL1 182
8.7.1 Image irradiance 182
8.7.2 Discussion 183
8.7.3 Focal spot-size 185
8.7.4 Discussion 186
8.7.5 Depth of focus 186
8.7.6 Spherical aberration 187
8.7.7 Acousto-optic resolution tests 187
8.7.8 Discussion 188
8.8 Overlay Lens LL1 190
8.8.1 Results 190
8.8.2 Discussion 191
8.9 Geodesic Lens GL1 192
8.9.1 Photographs 192
8.9.2 A computer-acquired image 192
8.9.3 Irradiance profiles 193
8.10 Substrate GL2 195
8.11 Focal Length of Lens SL1 197
8.12 Conclusions 199
References 201

CHAPTER 9 : CONCLUSIONS AND RECOMMENDATIONS FOR FUTURE WORK
9.1 Review of Thesis 204
9.2 Conclusions on Luneburg Lenses 210
9.3 Recommendations for Future Work 210
9.4 Conclusions on Geodesic Lenses 212
9.5 Recommendations for Future Work 213
9.6 Conclusions on Propagation Models 214
9.7 Recommendations for Future Work 215
9.8 Summary of original work 215a

PUBLICATIONS 216

CHAPTER ONE

INTRODUCTION

CHAPTER ONE INTRODUCTION

INHOMOGENEOUS LENS STRUCTURES FOR INTEGRATED OPTICS

Data-processing, interpreted as computing in its widest sense, is becoming progressively more complex [1]. Historically, many of the advances in processing have both driven and been driven by advances at the electronic component level. The phenomenon of 'silicon shrink', faster switching speeds and lower switching energies have triggered performance gains at the system level which have expanded data-processing horizons. There are signs, however, that the interdependence of electronics and data-processing systems at the most ambitious levels is coming to an end. There are two reasons:

1) the rate of advance in electronics at the component (ie chip-) level is slowing. Fundamental limits are being approached in terms of feature size and switching speeds

2) systems designers are increasingly hide-bound by von Neumann-oriented architectures which operate in sequential fashion. Processors have been essentially memory-driven rather than data-driven, preventing many operations from taking place in real-time. Bottlenecks are created which impede data flow and lead to interconnection difficulties with the real world.

There is a strong impetus to develop new types of computing engine which will overcome these problems. These computers will be required to operate in a massively parallel fashion and thus display much larger space-bandwidth products than present day electronic systems. Furthermore, an increase in speed (corresponding to an increase in time-bandwidth product) is also desirable. Pattern recognition, real-time spectral analysis, synthetic aperture radar, machine-vision and artificial intelligence are areas in which processors having these improved

characteristics could be of value. Purely electronic systems, arguably, are unlikely to prove suitable for their construction.

An alternative technology may thus be required. One which is suitable on several counts is based on optics. Light beams are inherently, and massively, parallel. Conceptually at least, millions of resolvable spots are available, each of which represents an independent signal-channel substantially free of interference or cross-talk. Because of the high frequency of light each channel can support an enormous temporal bandwidth. Beams can propagate through each other in many media at high power densities without interacting, while useful interactions within suitable non-linear materials can take place on a femto-second to pico-second timescale. The advances in communication using optical fibres are further reasons for developing all-optical processing, minimising the need for slow, power-costly photon-electron conversions.

Optics has been of considerable interest for performing analogue signal-processing for at least two decades, mainly as a result of the development of highly coherent sources,ie lasers. The operations of convolution, correlation, spectral analysis and imaging are relatively easy to perform in optics and these operations have formed the building blocks for powerful but highly specialised systems. More recently, attention has turned to the possible implementation of numerical (digital) optical processing as a means of obtaining systems of more general utility [2]. Classical optical elements such as lenses, mirrors, beam-splitters and prisms can be used in conjunction with acousto-optic and electro-optic interactions to perform algebraic operations, such as matrix-matrix and matrix-vector multiplications and additions. Analogue operations such as convolution can generally be configured as matrix-vector problems, so that a move towards numerical optical processing offers the likelihood of performance gains associated with digital processing (eg increased accuracy and reliability) in those areas.

It is evident that optics is an exciting approach to fast numerical processing implemented in a parallel fashion. The parallelism derives from the extension of a light wave-field over a wide region of three-dimensional space in comparison with the wavelength of the light. Somewhat paradoxically, there are advantages to be obtained in not using the full dimensionality offered by bulk optics. Reducing the dimensionality by a factor of one (and sometimes two) by confining the light to a planar wave-guide offers much greater control over several of the different processes involved. Guided-wave optics implemented on flat substrates is now well-established and is generally known (perhaps misleadingly) as integrated optics.

Some particular advantages quoted of integrated optics over bulk optics are, typically, decreased power consumption, convenient fabrication procedures, improved interaction efficiencies with other types of guided wave, longer interaction lengths owing to the non-spreading of the confined beam and gains in ruggedness, reliability, environmental immunity, small size and, particularly, the likelihood of compatibility with optical fibres. A major disadvantage is the loss of one transverse dimension leading to a much-reduced channel carrying capacity.

Caulfield [3] has pointed out that while the lateral spatial parallelism of bulk optical processors is not fully realized in integrated optics, a compensating form of discrete longitudinal parallelism is gained in which the incoming optical signal extending over one transverse dimension is modulated by time-varying and space-varying quantities such as presented by surface acoustic waves. Psaltis [4] makes some pertinent comments that may be quoted in full:

> ' The 2-D Fourier transforming lens is the primary mechanism that is used to make all the possible inter-connections in a 2-D optical processor. However, only a relatively small number of linear operations can be performed with a processor based on the 2-D Fourier transform. It is in fact

possible to compute a wider class of linear operations through combinations of imaging 1-D transforming (or integrating) lenses. Furthermore, greater flexibility exists in the design of such processors, ...'

Psaltis does not have integrated optics in mind here - his 'one-dimensionality' refers to a combination of non-guided-wave acousto-optic devices, light-source arrays and detectors, and bulk cylindrical lens elements - but the rationale is clear and a role for integrated optics seems possible.

The crucial part which the classical optical element, the lens, has to play in optical processing is stated emphatically in the paragraph quoted above. It is well-known that the optical diffraction process, in the Fraunhofer (or far-field) zone, is essentially the Fourier transformation process. On the surface of a sphere with centre situated in the plane of the diffracting aperture, transparency or object, the amplitude of the field is directly proportional to the Fourier transform of the aperture function [5]. The major problem involved in using Fraunhofer diffraction to obtain the Fourier transform is the need, in general, to use long observation distances. A thin lens can be used to overcome this problem, since the focusing action of the lens has the effect of compressing the distances involved. In the focal plane of a thin lens, the intensity of the diffraction field is proportional to the Fourier spectrum of the incident wave.

Lenses can be used for tasks other than Fourier transforming, such as beam collimation and expansion, spatial filtering, imaging and as summation/integration elements.

Real lenses fall short of the perfect optical imaging properties ascribed to ideal lenses. The errors introduced in imaging are quantifiable in terms of geometrical optics (less so in physical optics) and are called aberrations. The principal design tools available to lens-makers for correcting aberrations

have been, traditionally, the manipulation of surface-curvatures in lens elements based on spherical geometries. Combinations of these elements, sometimes including elements of different refractive index, have been used to minimise aberrations of a given type. It is not usually possible to minimise all aberrations simultaneously. Developments in machine tools enabled designers to introduce aspheric geometries readily into their designs, though aspheric lenses are still relatively rare, due to the not inconsiderable expense involved in their design and construction.

The refractive index of the material used in the fabrication of a single lens element was not considered a design variable and was almost invariably held constant. Major advances have occurred in this area however, largely due to the influence of the applied mathematician, Rudolf Karl Luneburg (1903-1949). His synthesis of the disciplines of physical optics, geometrical optics and electro-magnetism (that had previously led largely separate though related existences) enabled him to formulate single element designs incorporating variable refractive index functions. The crucial aspect of these lenses was that they were perfect; no correction by other elements was required, and 'perfection' (perfect imaging) was obtained over a solid angle of 360 degrees. In the 1950's and 1960's Luneburg lenses found some application in the microwave field, especially for high-fidelity radar-scanning and tracking purposes. The lenses were not extensively used in optics due to difficulties involved in construction.

The importance of these lenses for integrated optics became clear for two reasons: one being that the effective refractive index dispersion with waveguide thickness offered the possibility of simple construction, albeit at the expense, once again, of reduced dimensionality; and the other being that the scattering losses associated with integrated optical waveguides, especially at index discontinuities such as interfaces, precluded the use of several cascaded elements to achieve good image correction.

The principle of equivalent optical paths allows the construction in integrated optics of lenses analogous to Luneburg lenses which depend on the controlled geometrical deformation of the waveguide for successful operation. Such lenses are termed configurational or geodesic lenses. Geodesic lenses are strongly aspheric in general and are therefore more expensive to manufacture than Luneburg lenses. However advantages in terms of fabrication and performance over Luneburg lenses can be gained.

A multi-disciplinary approach is required in the investigation of these lenses, as is the case with much of integrated optics. The design and analysis of lenses involves considerable theoretical and computational effort; lens fabrication involves both theoretical and practical aspects of materials science, thin-film and vacuum technology and machining, polishing and masking problems; the characterization of fabricated lenses draws upon classical optical practices supplemented by new techniques unique to integrated optics; and incorporating the lenses into actual data-processing systems involves all of these again, together with conventional electronic engineering techniques, as well as other disciplines such as acousto-optics.

This thesis describes the synthesis of several of the above-mentioned disciplines in an effort to construct and operate two types of lens having common origins in the work of Luneburg. These lenses are likely to play major roles in the future development of integrated optics within the wider context of optical processing. In chapter two of the thesis the historical development of inhomogeneous lenses will be reviewed. The review will form the basis for an investigation into the design theory of inhomogeneous lenses for integrated optical applications. Two types of inhomogeneous lens will be identified which are closely related to one another; the thin-film Luneburg overlay lens and the geodesic surface-depression lens. The theory of formation of thin-film lenses using shadow masking techniques will be discussed in chapter three. Chapters four and five will both deal with the theoretical optical performance of inhomogeneous lenses.

Chapter four will analyse the performance of the lenses from a geometrical optics point of view, ray-tracing being used to quantify the aberrations that arise when the lens profiles obtained in the real world do not exactly correspond to the theoretically perfect profiles. Chapter five uses a numerical technique known as the beam-propagation method to investigate the characteristics of the lenses in terms of wave-optics. The technique has several advantages over conventional wave-optics methods. Diffraction problems, the effects of both small and large aberrations, off-axis incidence, beams of variable phase and amplitude, and propagation in inhomogeneous media are all readily dealt with by the method.

Chapter six will deal with the fabrication of planar waveguides and thin-film overlay lenses. Chapter seven will deal with the problem of fabricating uniform waveguides on the highly curved geodesic lenses investigated. Experimental measurements on several lenses will be reported in chapter eight. Finally, the various threads of the thesis will be pulled together in chapter nine and recommendations for future work will be proposed.

REFERENCES FOR CHAPTER ONE

[1] 'The State of Computing', COMPUTER Magazine, 17(10), October 1984

[2] Special issue on Optical Computing, Proc. IEEE, 72(7), July 1984

[3] Caulfield,H.J.,
Integrated optic pipelined optical signal processors, Wave Electronics, 4(4), 1983, 237-244

[4] Psaltis,D.,
see ref [2], pp 962-974

[5] Papoulis,A.,
'Systems and Transforms with Applications in Optics', McGraw-Hill, New York, 1968

CHAPTER TWO

INHOMOGENEOUS LENS DESIGN

CHAPTER TWO

THE DESIGN OF INHOMOGENEOUS LENSES

2.1 Introduction

Stigmatic imaging is feasible, conceptually, within the geometrical optics approximation [1]. An infinite number of rays emanating from a point source in object space may be made to pass through a single conjugate point in image space by an optical instrument using either reflection or refraction or a combination of both.

An imaging process may further be described as perfect if all the points in the object space are imaged in the above fashion with the result that the image points are geometrically mapped in one-to-one correspondence with the object points. Such a process is an abstraction and is unrealizable. Plane mirrors and holographic elements come reasonably close, however.

A more feasible and restricted aim is to image perfectly all the points on a sheet (surface) or even a line function. The demands made on the optical instrument are nevertheless still formidable. The use of the adjective 'perfect' in the remainder of this thesis will be intended to convey imaging in this restricted sense.

2.2 Examples

(a) The Cartesian Ovals

A class of lenses which performs stigmatic imaging has been known for over two hundred and fifty years and yet optical designers have found little use for them as they are of aspheric geometry. Descartes first stipulated the required geometry and, in consequence, the lenses are known as the Cartesian ovals [2]. It is interesting to note that the Cartesian ovals predate the first reference to Snell's law of refraction and consequently a

controversy (which remains unresolved) arose as to whom the law should be credited, with Fermat and Descartes being the principal protagonists.

The Cartesian ovals reduce to conic sections in the case of the imaging of a point situated on the optical axis at infinity. Workers in integrated optics only have two dimensions to contend with, generally, and it is reasonably obvious that generating a non-circular geometry in a plane should be much easier than generating aspherical shapes in bulk optics. Consequently the Cartesian ovals have been of some utility in integrated optics and both plano-hyperbolic [3] and elliptical [4] shapes have been utilised. The elliptical lens is remarkable in that it performs a positive focusing action while possessing a double-concave geometry. It manages to do this because the effective index within the lens is constrained to be lower than the effective index of the ambient waveguide. Such a situation is probably unheard of in bulk optics, and represents an added degree of freedom for design purposes.

(b) The sphere

Cartesian ovals are not capable, however, of performing perfect, geometrical imaging of extended lines or surfaces. A structure that is capable of doing this is the sphere, where the object points are those on the surface of the sphere and the rays are assumed to be confined to the surface. The rays emanating from those points travel along the great circles (geodesics) of the sphere and thus meet again at the diametrically opposite point. A hemispherical lens has been made to demonstrate this in integrated optics [5].

(c) Maxwell's fish-eye lens

An optical system with an index of refraction characterised by a rotationally symmetric function of the type:

$$n(r) = a/(b^2+r^2) \qquad (2.1)$$

where n(r) is the refractive index, r is the radial coordinate and a and b are constants, is known as a Maxwell's fish-eye system [1]. The fish-eye lens is able to image perfectly one sphere onto another without distortion.

(d) Luneburg lenses, and other inhomogeneous lenses

Refracting structures in which the optical rays follow curved rather than linear paths have aroused considerable interest over the past four decades, originally in the microwave field [6] and, more recently, in integrated optics [7]. The technologies required to produce the inhomogeneous lens structures which guide the rays are readily available in integrated optics, although the lenses of course operate only in two dimensions, as opposed to the three dimensions often envisaged by the original designers. Curved rays are not unknown in nature; the eyes of certain marine animals are known to produce this phenomenon [8]. Mirages and the propagation of radio waves in the upper atmosphere display similar characteristics. The interest in using inhomogeneous media derives from the possibilities of using such structures to yield perfect imaging.

R.K. Luneburg, in a classic but almost unobtainable text [9], dealt with the mathematical problems of inhomogeneous lenses in considerable depth. He formulated mathematical descriptions of a class of perfect inhomogeneous optical lenses now known as Luneburg lenses. These will be discussed in detail in the next section. Firstly, however, two distinct lines of possible technological development of inhomogeneous lenses will be identified.

2.3 Two Lines of Development

2.3.1 The principle of equivalence

The Luneburg lens and the Maxwell fish-eye lens are examples of lenses with variable refractive index distributions. The

physical boundaries of these lenses are generally spherical. The refractive index distributions can, however, be confined to a plane, in which case the boundaries of the lenses are described by circles.

Kunz [10] and Toraldo di Francia [11] discovered, independently, the possibility of finding a non-Euclidian (ie non-planar) two-dimensional space having the same optical properties as the planes in which inhomogeneous structures such as Maxwell fish-eye and Luneburg lenses resided. Mathematically, in accordance with Fermat's principle, the physical path followed by a ray between two points A and B is such as to make the variation in the optical path length zero:

$$\delta \int_A^B n dL = 0 \qquad (2.2)$$

n is the refractive index and dL is a line element along the path. Zero variation means that the optical path length along the ray is a local maximum or minimum (usually the latter). If the refractive index of the medium is constant then the variation in the physical path length for all the rays is also zero:

$$\delta \int_A^B dL = 0 \qquad (2.3)$$

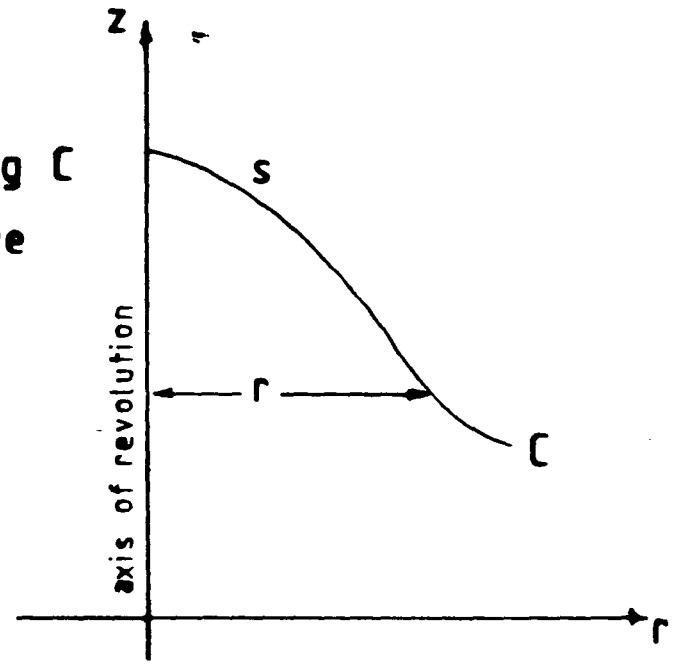
The rays, consequently, follow the geodesics of the surface.

In rotationally symmetric media, a surface of revolution is obtained by rotating a generating curve C around an axis of revolution z, as shown in Figure [2.1]. If s is taken as the arc-length along the curve, and (r,Θ) taken as the remaining cylindrical coordinates, the physical distance between neighbouring points on the curve C is given by:

$$dL^2 = ds^2 + r^2 d\Theta^2 \qquad (2.4)$$

s = arc length along C

C = generating curve



FIGURE[2.1] GENERATING CURVE OF A SURFACE OF REVOLUTION

The optical distance along these points is obtained by multiplying the physical distance by the index of refraction:

$$(ndL)^2 = n^2[ds^2 + r^2d\Theta^2] \qquad (2.5)$$

Equivalent optical surfaces are surfaces that identically preserve the relationship given by (2.5) without necessarily identically preserving the relationship given by (2.4). Put simply, equivalent optical surfaces have identical optical, though not necessarily physical, geometries.

A non-Euclidian two-dimensional analogue of Maxwell's fish-eye lens is the surface of the spherical lens described previously. Similarily a non-Euclidian lens equivalent to the Luneburg lens exists and is called the Rinehart lens [10]. Non-Euclidian two-dimensional lens surfaces may conveniently be called configurational lenses. In cases where the refractive index of the medium is constant, the lenses may be referred to as geodesic lenses.

Geodesic lenses are inhomogeneous in the sense that the profile curvature varies from point to point on the lens, in general. Lenses which have a point to point variation in refractive index are also inhomogeneous, and are called variable-index lenses for convenience. Two distinct paths of technological deveopment were available, therefore, for the designers of inhomogeneous lenses: the optical properties of the lens could be controlled by either continuously varying the physical geometry of the lens, or by continously varying the refractive index function.

The difficulties of physically obtaining aspheric structures, in bulk optics, are considerable; the problem of realizing a continuously varying radially-symmetric refractive index function is almost intractable. The dispersion of mode effective refractive index with waveguide thickness in guided-wave optics, however, provides an accessible means of realizing

variable-index lenses. Aspheric machining is a means available for producing the geodesic inhomogeneous lenses; the waveguide, whilst remaining of more-or-less constant depth, may be deformed so that the optical path of rays traversing the lens may be controlled in a manner appropriate for perfect imaging.

In this thesis, integrated optical lenses of both the geodesic and variable-index types will be discussed. It is worthwhile to trace the parallel development of these lenses from the pioneering work done by R.K. Luneburg, through the applications found for these lenses in microwaves, to the present-day interest in the lenses in the field of integrated optics.

2.3.2 Variable-refractive-index lenses

Luneburg [9] developed integral equations for variable refractive index lenses which were capable of imaging perfectly all the points on a sphere situated in homogeneous space on or outside the lens surface onto another conjugate sphere also situated outside the lens. In particular he found an analytical solution for a lens that could perfectly image the points on a sphere coincident with the outer surface of the lens itself onto a sphere situated at infinity. He had thus described a method of producing, from a point-source feed, a perfectly collimated beam.

Luneburg lenses which were capable of focusing a point at infinity to a point inside the lens itself were subsequently proposed by Brown [12] and Gutman [13]. The motivation of both Brown and Gutman in providing for a focus inside the lens was to increase the rapidity at which a point-source could mechanically scan the focal surface whilst maintaining the output beamwidth of the original design. Brown introduced the possibility of arbitrarily specifying the refractive index function of an outer shell from which the refractive index function of the perfectly focusing inner shell could then be derived. He also considered the problems of diffraction connected with the finite wavelength of the emerging radiation and proposed a tapering of the beam

amplitude to counteract these effects, a process now well-established in optics and elsewhere in wave systems and known as apodization.

Morgan [14] produced a general theory of Luneburg lenses which allowed for any combination of conjugate foci and a finite number of refractive index discontinuities within the lens. The refractive index function was specified by combinations of definite integrals, most of which could only be evaluated using numerical methods. Particular examples of solutions in terms of tabulated values were specified for lenses having focal lengths of between one and two times the lens radius.

These lenses subsequently found actual application in microwaves, either as collimating elements for point-sources or as wide-angle receiving antennae. One method of physically implementing the lenses involved approximating the required refractive index profile by constructing a number of concentric spherical shells, each shell having a different value of dielectric constant [6]. A lens of 18" (43cm) in diameter, composed of 10 shells made of styro-foam of different densities was shown to produce a wavefront phase-error in the exit pupil of less than one-tenth of a wavelength, a value much less than the commonly quoted quarter-wave limit for defect-free imaging.

In keeping with the philosophy first outlined by Miller [15], the new technology of integrated optics has borrowed extensively from microwaves. This has certainly been true for Luneburg lenses. The integrated optical designer is presented with a new degree of freedom compared with the bulk optics designer, though not compared with the microwave engineer, in having refractive index dispersion available as a tool. Ulrich and Martin [16] developed a two-dimensional geometrical optics theory which yielded ray and eikonal equations formally equivalent to those of classical optics [1]. Van Duzer [17] worked along similar lines in the related field of surface-guided acoustic waves. Consequently the design of surface guided-wave components was made amenable to ray methods.

Shubert and Harris [18] pointed out the favourable possibilities of integrated optical signal-processing in comparison with bulk optics, in spite of the inherent reduction in dimensionality, or parallelism, and went on to give a modal analysis of light propagation in multi-layered thin films that gave full scope for utilising the achievable variation in the surface-wave eigenvalues; effectively one had the possibility of control over refractive index by simply changing material parameters and varying deposited film thicknesses. Whilst not producing Luneburg lenses in the strictest sense, these workers were among the first to recognize the potential of controlled inhomegeneity in integrated optical lenses. Lenses with f-numbers as low as f/2 were considered achievable, and various layer configurations were envisaged, including lenses over- and under-laid with respect to the ambient waveguide. The use of cladding layers to reduce losses was also envisaged. A non-circular lens structure having a Gaussian effective refractive-index distribution in the propagation direction and a parabolic transverse distribution was presented, and motor-driven shadow masks for use in shaping thin films were stated to be capable of producing film lens and prism layers of any desired thickness profile.

Suematsu et al [19] dealt with the focusing properties of a thin-film lenslike light guide having an approximately parabolic transverse film thickness profile. They proposed using the lens for phase-matching in a thin-film optical second-harmonic generator. They reported use of a shadow-masking technique in order to modify the thickness profile, with the mask situated in the space between a sputtered source and the target substrate. The lenses so produced were conceptually similar to gas lenses and self-focusing (SELFOC) optical fibres.

Zernike [20] was the first worker (to the author's knowledge) to explicitly advocate the use of the lens profiles formulated by Luneburg in integrated optics. He drew on the work of Tien et al [21], on tapering the regions interfacing two areas

of different effective refractive index, to realize the required profile. A shadow-masking geometry was used to produce a lens having an overall diameter of 11.5mm. Niobium pentoxide (Nb_2O_5) was deposited in a controlled fashion onto planar waveguides on glass substrates. The shadow mask was bevelled in order to produce a slow tapering region between the outer waveguide and the central part of the lens. Little mention was made of the principles applied in either lens or mask design, and the presumption may perhaps be made that these were done on an ad hoc basis. Nevertheless the results obtained were reasonably good. A parallel bundle of rays filling almost the entire lens aperture focused onto a point near the lens edge and then diverged at an angle of 110 degrees. A small amount of spherical aberration was manifested. Another lens, focusing at a point outside the lens surface produced a focal spot diameter of less than 30um from an input beamwidth of 5mm.

Two papers by Anderson et al [22,7] placed the Luneburg lens squarely at the forefront of integrated optical technology. Lenses were to be incorporated as both collimating and spatial Fourier transforming elements in the integrated optical spectrum analyzer (IOSA). The thin film lens was felt by Anderson and his colleagues to be the most viable planar lens for low substrate refractive index applications such as those based on SiO_2-on-Si technology. For higher index substrates such as lithium niobate ($LiNbO_3$) it was felt that the configurational or geodesic type of lens would be more appropriate since few waveguide materials were available which could yield the refractive index changes necessary for the low f-number applications required. The fine resolution required of the IOSA demanded diffraction-limited, aberration-free performance from the lenses, thus also requiring precise control of the lens profiles. The experimental results quoted by these workers suffer from a lack of correspondence between figure captions and text, but the conclusions of the study were that thin-film Luneburg lenses could be produced which exhibited only a small degree of spherical aberration.

W.H. Southwell [23] made a significant contribution to both

design and analysis of thin-film Luneburg lenses by approximating the hitherto cumbersome transcendental expressions for the refractive index functions with simpler analytical expressions which were readily incorporated in computations. The methods presented by Southwell were utilised extensively by the present author and will be reported in some detail.

The most impressive achievments so far in the field of thin-film Luneburg lenses have come from Yao et al [24,25]. Computer-modelled shadow-masking geometries were optimised to achieve as close a synthesis as possible to the required thickness profiles. The results were excellent, near diffraction-limited focusing properties being obtained.

Recent developments in the field of overlay Luneburg lenses are reported in references [26,27,28,29]. Of particular interest is the work carried out by Busch et al [28] which parallels that of the author to some extent. Comparisons will be made between the author's findings and those reported in reference [28].

2.3.3 Configurational/geodesic lenses

Rinehart [30] derived a geodesic analogue of the classical Luneburg lens, which transformed the planar, variable refractive index Luneburg lens into a surface of revolution of constant refractive index which reproduced the optical properties of the planar system. Rinehart's motivation was to solve the problem of rapid scanning for radar antennae over a large angular domain.

Kunz [10] and Toraldo di Francia [11] extended the work of Rinehart and Luneburg by demonstrating that a family of configurational lenses of equivalent optical properties to the Luneburg lens existed. Table [2.1], taken from Kunz's paper, illustrates some of these possible equivalent optical systems, expressed in terms of the rotationally symmetric geometry of Figure [2.1] . Functions describing the physical geometry of the generating curve are tabulated, as well as functions describing the required refractive index distributions. It may be observed

TABLE [2.1] Lenses theoretically equivalent to the Luneburg lens (after KUNZ [10])

SURFACE	GENERATING CURVE	INDEX REQUIRED *	PERMISSIBLE RANGE
1. Plane (Luneburg lens itself)	$s = r$	$[2 - r^2]^{1/2}$	$0 < r < 1$
2. Cone	$s = kr$	$r^{k-1}[2 - r^{2k}]^{1/2}$	$(0 < r < 1$ if $k > 0)$ $(r > 1$ if $k < 0)$
3. Cylinder	$r = 1$	$e^{-s}[2 - e^{-2s}]^{1/2}$	$s > 0$
4. Rinehart surface	$s = 1/2(r + \arcsin(r))$	1	$0 < r < 1$
5. Generalized Rinehart surface	$s = (1-k)r + k \arcsin(r)$	$\dfrac{2[1 + (1-r^2)^{1/2}]^{2k} - r^2}{[1 + (1-r^2)^{1/2}]^{2k}}^{1/2}$	$0 < r < 1$
6. Hemisphere	$s = \arcsin(r)$	$\dfrac{2[1 + (1-r^2)^{1/2}]^2 - r^2]^{1/2}}{[1 + (1-r^2)^{1/2}]^2}$	$0 < r < 1$

* normalized to n=1 outside lens

that the Rinehart surface having a constant refractive index of 1 constitutes a geodesic system[1] (the functions are tabulated for a system with an ambient refractive index of 1).

Rudduck et al [31] developed non-planar lens geometries equivalent to the generalized Luneburg lenses of Morgan. Geodesic versions were implemented experimentally and excellent collimated radiation patterns were observed in the image space arising from a point-source feed situated at the focus.

Van Duzer [17] described configurational surface-depression and surface-protrusion lenses that could be used for focusing surface acoustic waves. Subsequently, Righini et al [32] gave formulae for producing aberration-free geodesic lenses analogous to the Rinehart lens for integrated optical purposes. Several types of lenses were constructed on glass substrates and tested successfully. Spiller and Harper [33] combined overlay films and spherical geodesic lenses in a single lens element, to minimise third-order and fifth-order spherical aberrations. Vahey and Wood [34] proposed that geodesic lenses be produced with aspheric geometries, again to minimise aberrations. A great deal of work has subsequently been carried out on geodesic lenses [35-41]. The particular interest in the present work is centred on the general solution to the problem of perfect geodesic lenses presented by Sottini et al [42]. A design based on their general solution was developed and constructed by Doughty et al [43,44] and results will be reported in this work.

2.4 Design of the Thin-film Luneburg Lens

2.4.1 The generalized solution given by Luneburg

The motivation for Luneburg's interest in the type of lens now named after him arose from a problem with the Maxwell Fish-eye lens. The Fish-eye lens is a conceptually perfect instrument within the geometrical optics approximation. It forms a precise point-to-point image of one sphere onto another. Both image and object spheres are located in regions where the refractive index

1) In general, a configurational lens has a variable refract index function and a variable, non-Euclidian geometry. A geode lens is a configurational lens with a constant refractive inde

varies, a somewhat undesirable fact. Luneburg inquired whether a refractive index function could not be found for the lens which allowed the conjugate spheres to reside in uniform, homogeneous space. Luneburg found appropriate solutions and these are entirely sufficient to describe the refractive-index profiles of the thin-film Luneburg lenses investigated in this work.

The geometry of the situation is shown in Figure [2.2] . Only the case of both conjugate foci external to the lens itself is considered. The refractive index of the medium outside the lens is normalized to 1, as is the radius of the lens. The object sphere is situated at a radial distance r_0 from the lens centre, with the image sphere situated at a distance r_1. The polar coordinates (r^*,Θ^*) define the point of closest approach of a ray to the origin, where the ray is uniquely defined by an initial direction α and position (r_o, π). The refractive index of the lens sphere is given by the spherically symmetric function n(r). It will be shown in chapter four, in connection with ray-tracing, that rays in a spherically symmetric medium are plane curves which lie in planes through the origin, such that:

$$\Theta = K\int \frac{dr}{r\sqrt{(n^2(r)r^2 - K^2)}} \qquad (2.6)$$

where:

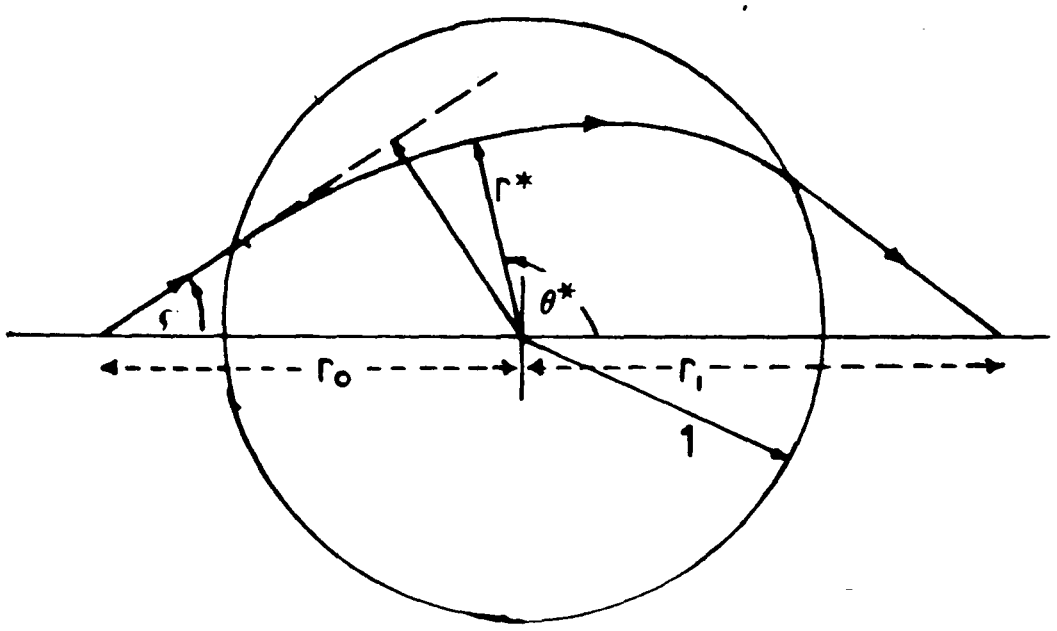
$$K = r_o \sin\alpha \qquad (2.7)$$

is a quantity called the ray constant, for obvious reasons.

A convenient variable may be defined which often arises in the theory of radially-symmetric media:

$$\rho(r) = rn(r) \qquad (2.8)$$

The function $\rho(r)$ is assumed to be a monotonically increasing function of r, which excludes the possibility of more than one



FIGURE[2.2] GEOMETRY OF GENERALIZED LUNEBURG LENS : CASE OF TWO EXTERNAL FOCI

extreme value of r^*. It is easily shown, using equation (2.6), that the equation of the light ray, after it has travelled beyond the point of closest approach to the origin, is given by:

$$\Theta = \pi + K\int_{r_o}^{r^*} \frac{dr}{r\sqrt{(\rho^2 - K^2)}} - K\int_{r^*}^{r} \frac{dr}{r\sqrt{(\rho^2 - K^2)}} \qquad (2.9)$$

so that the intersection point of the ray with the axis, $(r_1,0)$, is given by:

$$0 = \pi + K\int_{r_o}^{r^*} \frac{dr}{r\sqrt{(\rho^2 - K^2)}} - K\int_{r^*}^{r_1} \frac{dr}{r\sqrt{(\rho^2 - K^2)}} \qquad (2.10)$$

If the refractive index function is a given quantity, then equation (2.10) determines the intersection distance r_1 of a given ray as a function of K, ie, as a function of the initial direction of the ray. If r_1 is required to be constant, however, to yield perfect focusing, equation (2.10) represents an integral equation for the function $\rho(r)$ for $r < 1$. For $r \geq 1$, $\rho = r$. Now, since:

$$K\int \frac{dr}{r\sqrt{(r^2 - K^2)}} = -\arcsin(K/r) \qquad (2.11)$$

the integration of equation (2.10) may easily be carried out for the regions outside the lens proper, ie:

$$0 = \pi + K\int_{r_o}^{1} - K\int_{1}^{r^*} - K\int_{r^*}^{1} - K\int_{1}^{r_1}$$

$$= \pi + K\int_{r_o}^{1} - 2K\int_{r^*}^{1} - K\int_{1}^{r_1} \qquad (2.12)$$

where the integrands have been omitted, for convenience. Consequently, the condition:

$$K\int_{r^*}^{1} \frac{dr}{r\sqrt{(\rho^2 - K^2)}} = f(K) \qquad (2.13)$$

is obtained, where:

$$f(K) = \frac{1}{2}\left\{\pi + \arcsin(K/r_1) + \arcsin(K/r_o) - 2\arcsin(K)\right\} \qquad (2.14)$$

The variable:

$$\Omega(\rho) = -\log r(\rho) \qquad (2.15)$$

may be introduced to enable equation (2.13) to be solved. Equation (2.13) then becomes:

$$-K\int_{K}^{1} \frac{d\Omega(\rho)}{\sqrt{(\rho^2 - K^2)}} = f(K) \qquad (2.16)$$

which is an integral equation of Abel's type.

It can be shown, [9], that the following theorem is valid:

Theorem

if the function f(K) is defined by the integral:

$$f(K) = -K\int_{K}^{\lambda} \frac{d\Omega(\rho)}{\sqrt{(\rho^2 - K^2)}} \qquad (2.16)$$

in the interval $0 \leq K \leq \lambda$, then $\Omega(\rho)$ is determined by the integral:

$$\Omega(\rho) - \Omega(\lambda) = \frac{2}{\pi}\int_{\rho}^{\lambda} \frac{f(K)}{\sqrt{(K^2 - \rho^2)}} dK \qquad (2.17)$$

Using equation (2.17) it can be shown that the solution of equation (2.16) is given by:

$$\Omega(\rho) - \Omega(1)$$

$$= -\log(\rho) + \frac{1}{\pi}\int_{\rho}^{1} \left\{ \frac{(\arcsin(K/r_1) + \arcsin(K/r_o))}{\sqrt{(K^2 - \rho^2)}} \right\} dK$$

(2.18)

If the function:

$$w(\rho,s) = \frac{1}{\pi}\int_{\rho}^{1} \frac{\arcsin(x/s)dx}{\sqrt{(x^2-\rho^2)}} \qquad (2.19)$$

is defined, then:

$$\log(\rho) - \log(r) = \log(\rho/r) = w(\rho,r_o) + w(\rho,r_1)$$

(2.20)

since $\Omega(\rho) = -\log r$ and $\Omega(1) = 0$.

Finally, since $\rho = nr$, the required parametric equation of the refractive index profile of the lens is obtained:

$$n(r) = \exp\{w(\rho,r_0) + w(\rho,r_1)\} \qquad (2.21)$$

For the case of one sphere situated at infinity, $w(\rho,\infty) = 0$ and (2.21) becomes:

$$n(\rho,s) = \exp\{w(\rho,s)\} \qquad (2.22)$$

where s now describes the focal length of the sphere situated at the finite distance. This case will be dealt with exclusively henceforth.

It can be seen that the desired refractive index profile $n(r) = \rho(r)/r$ lies embedded in equation (2.22) due to the presence of the variable ρ in the integral equation (2.19). Equation (2.22) is, therefore, a transcendental relation. An analytical solution is available for the integral equation (2.19), for the case of an image sphere having a radius coincident with that of the lens. In such a case the required refractive index profile becomes:

$$n(r) = \sqrt{(2-r^2)} \qquad (2.23)$$

Lenses possessing such profiles are commonly called 'classical' Luneburg lenses while all other index profiles arising from different combinations of conjugate foci fall within the category of 'generalized' Luneburg lenses. For generalized Luneburg lenses, numerical methods must be used to evaluate the integral in equation (2.19). Southwell [23] developed a procedure for the numerical calculation which was adopted by the author.

2.4.2 Southwell's method

A change of variable, $y = x - \rho$, is first introduced in equation (2.19). The equation then becomes:

$$w(\rho,s) = \frac{1}{\pi}\int_0^{1-\rho} \frac{\arcsin[(y+\rho)/s]}{(y+2\rho)^{1/2}y^{1/2}}\,dy \qquad (2.24)$$

The $y^{-1/2}$ singularity is dealt with by performing an integration by parts. The resulting equation, which is not specified correctly in reference [23], becomes:

$$w(\rho,s) = \frac{1}{\pi}\left\{ 2\arcsin(1/s)\frac{(1-\rho)^{1/2}}{(1+\rho)^{1/2}} - 2\int_0^{1-\rho}\left[(s^2-(y+\rho)^2)^{-1/2} - \frac{\arcsin[(y+\rho)/s]}{2(y+2\rho)}\right]\frac{y^{1/2}dy}{(y+2\rho)^{1/2}}\right\} \qquad (2.25)$$

Equation (2.25), despite its apparent complexity, is in fact well-behaved and easily evaluated using numerical methods. Southwell divided the interval of integration in equation (2.25) into four regions and then performed a 32-point Gaussian quadrature on each . The present author found that a 64-point Gaussian quadrature over the complete interval was sufficient to obtain agreement to 1 part in 10^9 with Southwell's results.

2.4.3 An analytical approximation to $w(\rho,s)$

Although the integral in equation (2.25) can be numerically evaluated as described above, the desired refractive index distribution n(r) is still embedded according to equation (2.22). It is desirable, before going on to extract the required distribution, to derive an analytical approximation to $w(\rho,s)$. Such a procedure has several advantages:

(i) An analytical approximation (eg,a polynomial fit) to equation (2.25) can be constructed from a knowledge of the function at various points. The numerical integration need only be carried out at a discrete number of points in order to describe the function accurately over the complete interval.

(ii) The procedure to find the refractive index is a root-finding procedure. It is computationally more efficient to deal with an analytical approximation such as a polynomial fit than to have to perform a repetitive numerical integration.

(iii) The analytical approximation is easily incorporated into ray-tracing algorithms which help to assess the lens performance.

The behaviour of the function at limiting values can be investigated to give clues about the appropriate form of analytical representation. As ρ tends to 1 (close to the lens edge) the upper limit of integration in equation (2.25) tends to 0. The integrand of (2.24) then behaves as $y^{1/2}$ so that the integral portion of (2.25) behaves as $(1-\rho)^{3/2}$. This term is of higher order than the first term on the right hand side of (2.25). Hence as ρ tends to 1, equation (2.25) behaves as:

$$w(\rho,s) = \frac{2}{\pi}^{1/2} \arcsin(1/s)\,(1-\rho)^{1/2} \qquad (2.26)$$

It has already been observed that that the next-order term will have a $(1-\rho)^{3/2}$ behaviour, so that the following polynomial representation suggests itself:

$$w(\rho,s) = P_1(1-\rho)^{1/2} + P_2(1-\rho)^{3/2} + P_3(1-\rho)^{5/2} + P_4(1-\rho)^{7/2} + P_5(1-\rho)^{9/2} \qquad (2.27)$$

where the P's are the polynomial coefficients which remain to be determined. From (2.26):

$$P_1 = \frac{2}{\pi}^{1/2} \arcsin(1/s) \qquad (2.28)$$

Further constraints on the behaviour of the function (eg, at the limiting value as ρ tends to 0) may be incorporated. However, with the information thus far obtained it is possible to fit a

curve such as that specified in (2.27) to data obtained from the numerical calculation of equation (2.25) in order to evaluate the remaining coefficients.

The P coefficients obtained by the author for an s=2 lens are tabulated in Table [2.2], together with the published results of Southwell. The coefficients differ slightly in magnitude, though not in sign. The sums of squared errors over the complete aperture for lenses having various s-numbers are compared with Southwell's results in Table [2.3]. The values obtained by the author were slightly greater than those obtained by Southwell, but are still very small.

2.4.4 Refractive index profiles of Luneburg lenses

Newton's root-finding method is used to calculate the refractive index n(r) at a given radial distance r from the polynomial (2.27). From equation (2.22) a function F is defined as:

$$F(n) = \exp[w(\rho,s)] - n \qquad (2.29)$$

The value of n(r) which makes this function 'sufficiently small' is taken as the required refractive index of the lens at the point r. In practice, 'sufficiently small' may be taken as low as 10^{-8} without significantly slowing the calculation or affecting the convergence of the solution.

2.4.5 The thin-film overlay Luneburg lens

So far, the method outlined here to obtain Luneburg lens profiles can be applied to any medium of rotational symmetry in either two or three dimensions. Attention will now be restricted to applying the method in integrated optics, in which case a circularly symmetric profile is being considered.

TABLE [2.2] Coefficients of polynomial fit to $w(\rho,s)$ curve.

s=2 lens

COEFFICIENT	THIS WORK	SOUTHWELL
P1	0.235 699 486	0.235 687 835
P2	−0.074 917 543	−0.074 750 036
P3	0.007 348 242	0.006 728 945
P4	−0.005 984 959	−0.005 144 471
P5	−0.000 609 453	−0.000 989 300

TABLE [2.3] Residual sum of squared errors derived from polynomial fit to $w(\rho,s)$ curve.

s-number	THIS WORK	SOUTHWELL
s=2	4.8154×10^{-11}	3.9101×10^{-11}
s=3	8.3664×10^{-12}	5.8134×10^{-12}
s=5	2.5502×10^{-12}	1.7893×10^{-12}
s=9	7.6218×10^{-13}	5.3152×10^{-13}

A variation of thickness in a thin-film optical waveguide causes a dispersion in the 'effective' refractive index of the waveguide for a given mode [45]. The effective index is defined by:

$$n_e = \beta_m / k \qquad (2.30)$$

where β_m is the modal propagation constant and k is the free-space wave-number.

The situation envisaged for obtaining a refractive index variation such as that defined by equations (2.22) and (2.27) is shown in schematic form in Figure [2.3] . A substrate with material refractive index n_4 supports an ambient planar waveguide of material refractive index n_3, with $n_3 > n_4$. A material of high refractive index n_2 is used to form the lens.

A cladding layer of index n_1 is shown as air in the Figure, although other materials could be used.

In a sense, the thickness variation could be taken to mean that this lens should come into the configurational category. However, the variation is so small that any path length changes arising directly from the physical shape are negligible. The physical shape may be assumed to affect the optical path only in an indirect fashion through the variation of the effective index.

The depth, d, of the ambient waveguide is assumed to be constant. The thickness, t, of the overlay material is tailored to achieve a variation in effective index corresponding to a Luneburg lens design.

An assumption is made at this point. The materials used in the thin-film Luneburg lens are considered to be stepwise-homogeneous in the y-direction, perpendicular to the waveguide plane. Such an assumption is not always strictly correct; many waveguides have been fabricated which have a non-uniform

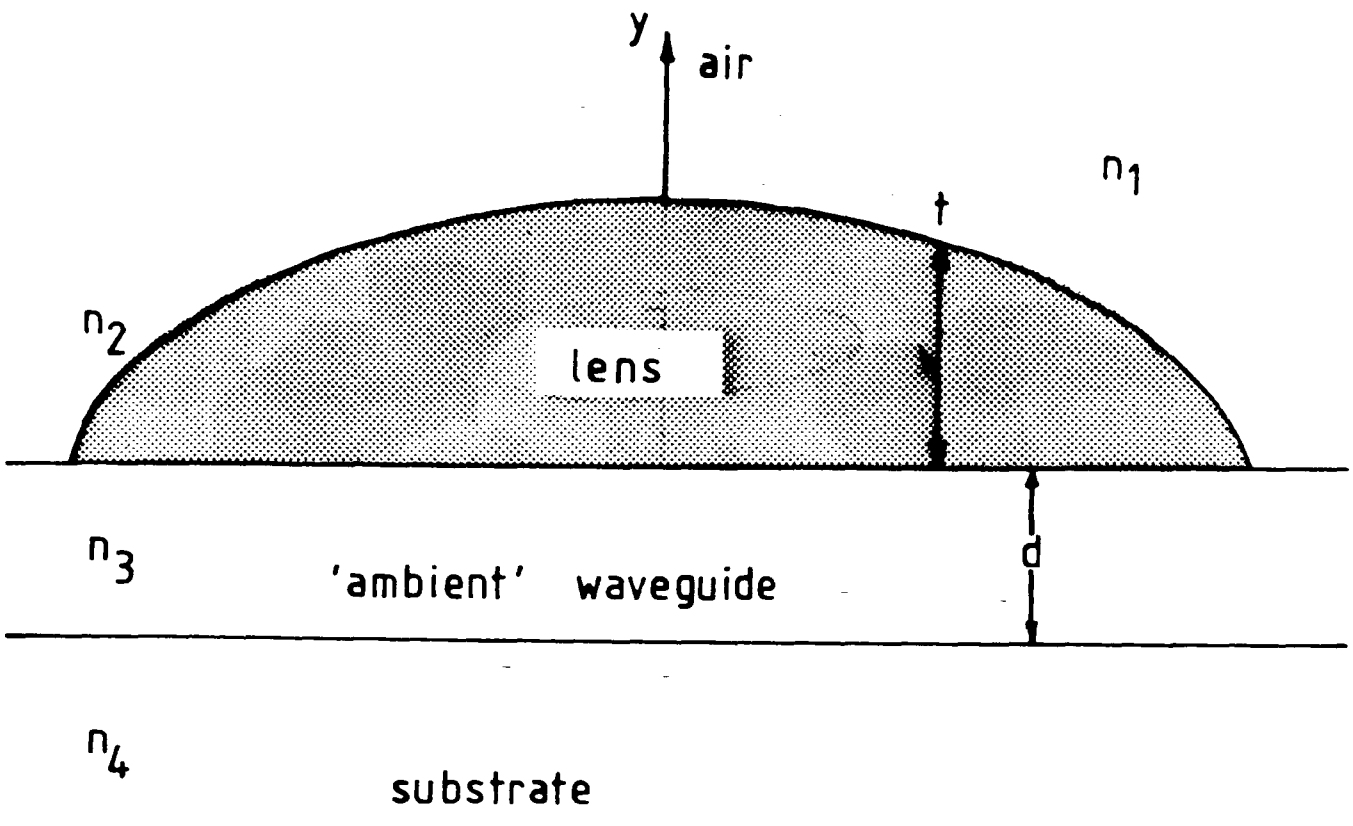


FIGURE [2,3] CROSS-SECTION OF FOUR-LAYER STRUCTURE FOR REALIZING A THIN-FILM LUNEBURG LENS

refractive index profile. Ti-indiffused waveguides are an example. Furthermore, the density of deposited thin films can vary with increasing deposition thickness. The electro-magnetic theory governing modal propagation in these cases is more complicated and an assumption of stepwise homogeneity greatly simplifies the calculations. The index difference obtained between ambient waveguides and $LiNbO_3$ substrates in the Ti-indiffusion process is very much smaller than the index difference between the overlay material, As_2S_3, and the ambient waveguide. It can be shown that more complicated models of the ambient waveguide in such a case lead to no substantial differences in the profile calculations [46].

The multi-layer modal propagation theory of Shubert and Harris [18] may be used to derive film thickness profiles from the calculated refractive index profiles. Attention will be restricted to TE modes. The theory governing the case of TM modes is only slightly more complicated.

The relationship between the effective refractive index n_e for TE modes and the layer dimensions and material indices shown in Figure [2.3] is given by:

$$-h/p = \frac{[1-(v/q)\tan(vkd)]\tan(hkt) \pm (h/v)[(v/q)+\tan(vkd)]}{[1-(v/q)\tan(vkd)] - (h/v)[(v/q)+\tan(vkd)]\tan(hkt)}$$

$$\text{for } n_e < n_3$$

or:

$$-h/p = \frac{[(q+v)\exp(2vkd)+q-v]\tan(hkt)+(h/v)[(q+v)\exp(2vkd)-q+v]}{[(q+v)\exp(2vkd)+q-v] - (h/v)\tan(hkt)[(q+v)\exp(2vkd)-q+v]}$$

$$\text{for } n_e \geqslant n_3 \qquad (2.31)$$

where:

$$k = 2\pi/\lambda$$

$$p = (n_e^2 - n_1^2)^{1/2}$$
$$h = (n_2^2 - n_e^2)^{1/2}$$
$$q = (n_e^2 - n_4^2)^{1/2}$$
and $$v = (n_3^2 - n_e^2)^{1/2} \quad \text{for } n_e < n_3$$
$$= (n_e^2 - n_3^2)^{1/2} \quad \text{for } n_e \geqslant n_3$$

(2.32)

The relationship between n_e and t as defined in (2.31) is transcendental; an explicit solution for n_e for a given thickness is not available. It is possible, however, to obtain an explicit solution for t for a given n_e:

$$kht = \tan^{-1}\left(\frac{h[1-(v/q)\tan(vkd)]+(ph/v)[(v/q)+\tan(vkd)]}{(h^2/v)[(v/q)+\tan(vkd)]-p[1-(v/q)\tan(vkd)]}\right)$$

$$+ m\pi \quad \text{for } n_e < n_3$$

$$kht = \tan^{-1}\frac{h[(q+v)\exp(2vkd)+q-v]+(ph/v)[(q+v)\exp(2vkd)-q+v]}{(h^2/v)[(q+v)\exp(2vkd)-q+v]-p[(q+v)\exp(2vkd)+q-v]}$$

$$+ m\pi \quad \text{for } n_e \geqslant n_3$$

(2.33)

The above expressions differ slightly from the ones quoted by Southwell, which are incorrect as published.

Plots of the dispersion curves for the first seven TE modes of a four-layer structure are shown in Figure [2.4]. The substrate refractive index is 2.2025 which is representative for the case of TE modes propagating in the crystallographic x-direction in y-cut $LiNbO_3$ at a free-space wavelength of 0.6328um. The waveguide is assumed to be a step-index guide although its refractive index, 2.2060, is actually typical of a waveguide formed by Ti-indiffusion (see chapter six of this thesis). The overlay material is a chalcogenide film, As_2S_3, which has a high material refractive index of 2.6 . The depth d of the ambient waveguide is assumed to be 3.5um, and the thickness t of the overlayer varies between 0.0 and 1.6um.

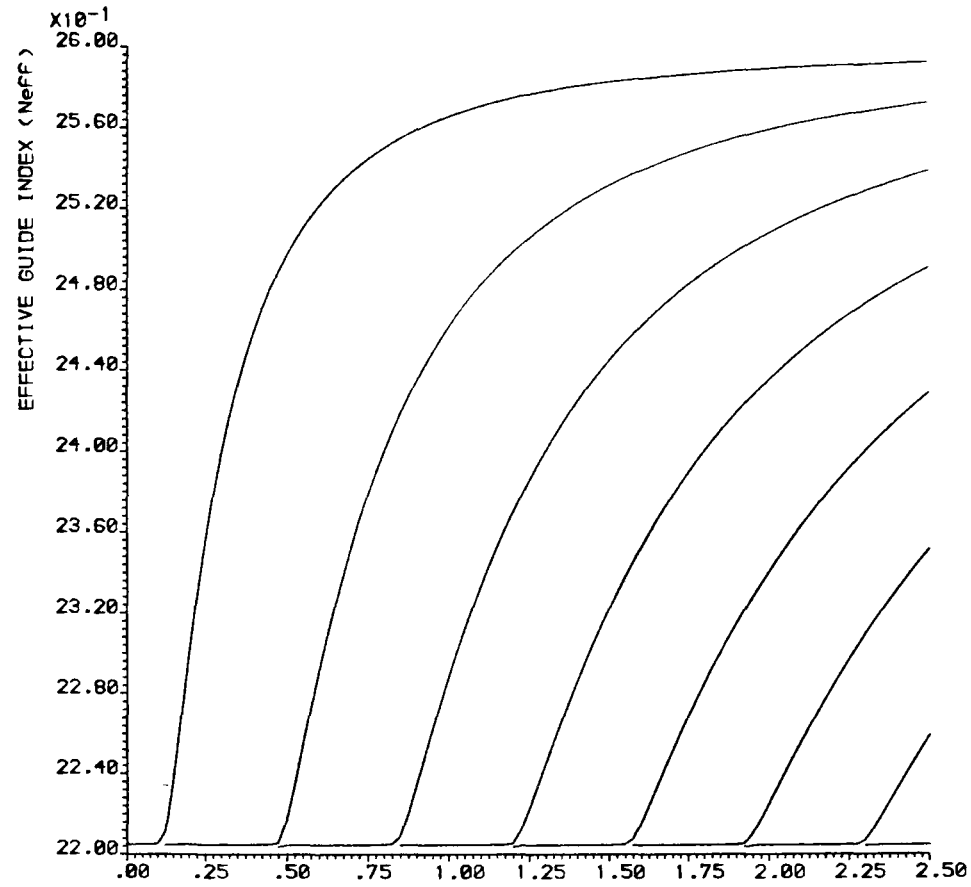


FIGURE [2.4]
DISPERSION CURVES (TE MODES) FOR A FOUR LAYER WAVEGUIDE

WAVEGUIDE PARAMETERS :

SUPERSTRATE - AIR; Na=1

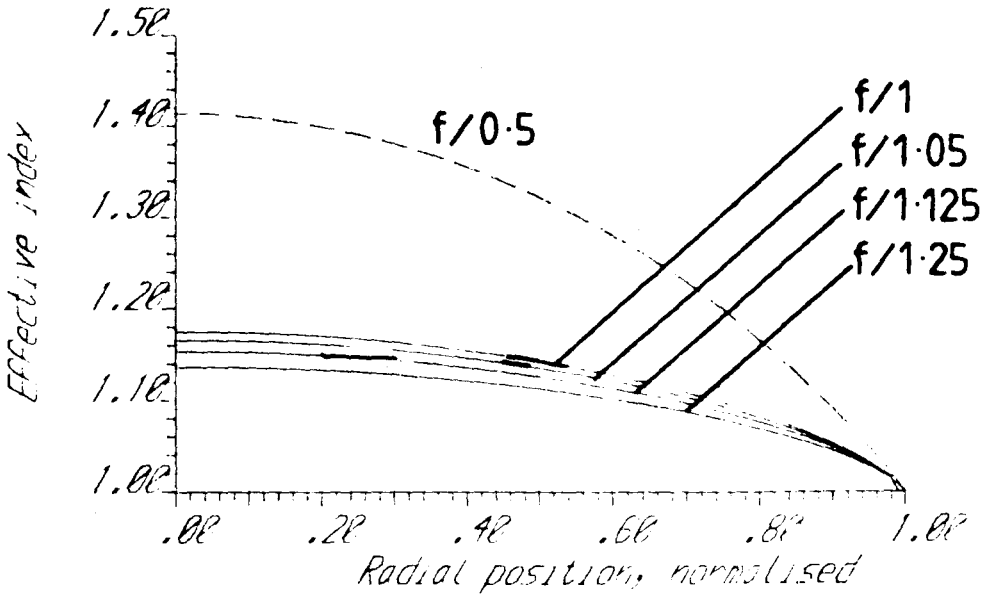
OVERLAYER - CHALCOGENIDE FILM; Nc=2.6
THICKNESS=VARIED

WAVEGUIDE - Y-CUT Ti-diffused LiNbO3 WAEGUIDE; Nw=2.2060(MAX.)
STEP INDEX PROFILE; D=5.53(WAVELENGTHS)

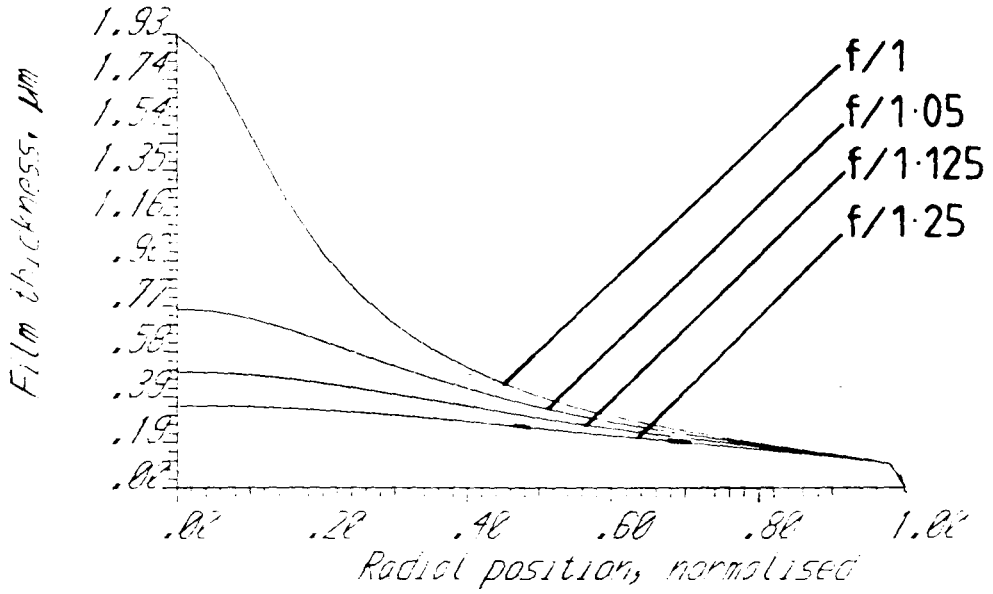
SUBSTRATE - LiNbO3; Ns=2.2025

Equation (2.33) may be used to calculate the thickness profiles required to realize thin-film Luneburg lenses . Index and thickness profiles are shown in Figures [2.5] and [2.6] for lenses with f-numbers of f/1, f/1.05, f/1.125, f/1.25, f/1.5, f/2, f/2.5, f/3 .

The lowest f-number that can be realised depends on the maximum change in effective refractive index that can be realised in the four-layer system. If the maximum index change is taken as the difference between bulk values of the substrate and overlay film indices, then a difference of approximately 0.4 is available for the As_2S_3/Ti:$LiNbO_3$ combination at a wavelength of 633nm. This index difference allows a lens of f/0.95 to be realized, theoretically. It is of interest to observe how the film thickness required at the centre of the lens depends on the desired f-number. The relationship is plotted in Figure [2.7]. An asymptotic behaviour is observed. For f-numbers in the the region of f/1, the centre thickness changes rapidly as a function of the required f-number. The thickness changes very slowly as a function of f-number beyond f/4. The significance of this behaviour for fabrication purposes may be illustrated by considering specific numerical values. An f/1 lens requires a centre overlay thickness of 1928nm, whereas an f/1.05 lens requires a centre thickness of 762nm. Thus a change in thickness of 1166nm only results in a 5% change in the focal length. The focal length is, therefore, not very sensitive to thickness variations for low values of f-number. However, the sensitivity greatly increases at larger values of f-number. An f/5 lens requires a centre film thickness of 116nm whereas an f/5.5 lens requires a centre thickness of 113nm. Thus a 10% change in focal length occurs for a change in centre thickness of only 3.0nm . Clearly lenses of even moderately large f-number would be difficult to fabricate such that the focal length could be controlled accurately. It would seem that lenses of very low f-number would be better candidates for fabrication. However, a re-inspection of Figure [2.5] shows that lenses with very low f-numbers display much more complicated overall profile behaviour than do lenses with larger f-numbers. Lenses with f-numbers f/1



(a)



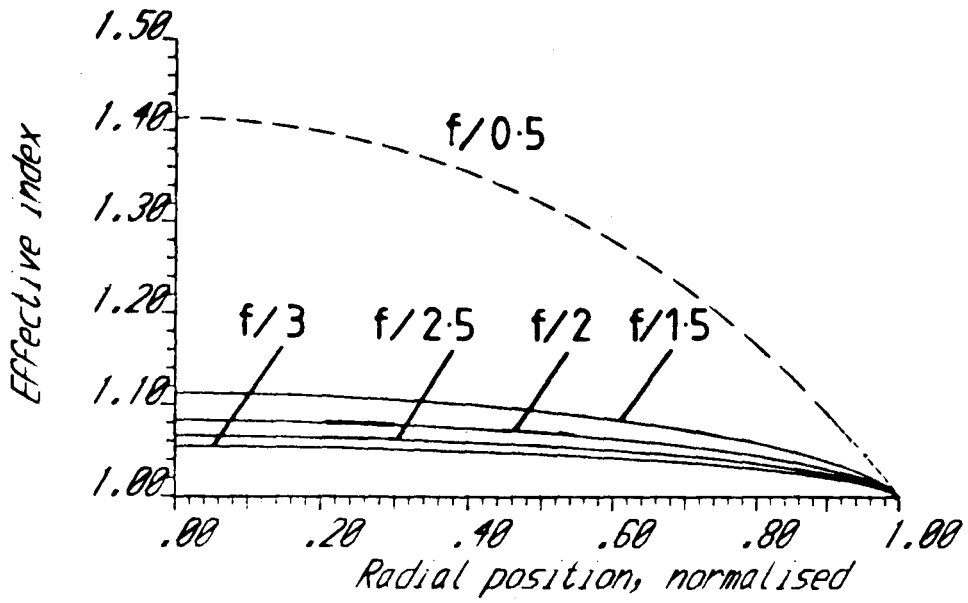
(b)

FIGURE [2,5] *Profiles for LUNEBURG OVERLAY LENSES*

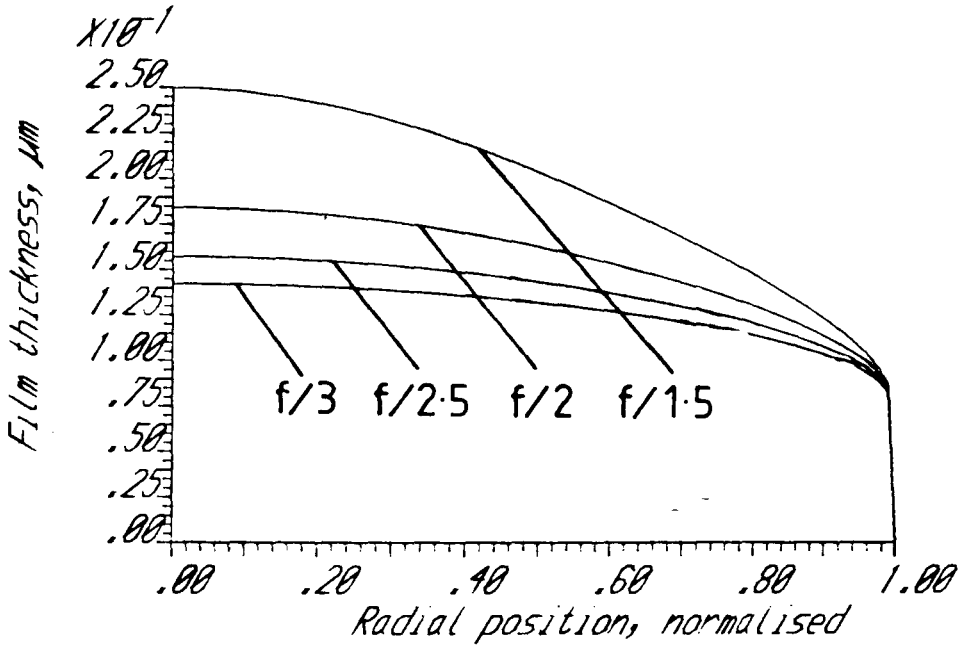
(a) effective index profiles, including 'classical' lens
(b) overlay film thickness profiles for synthesis of (a)

As2S3 film on Ti-diffused LiNbO3
y-cut, x-propagating

Film index= 2.6 Substrate index= 2.2028
Outer waveguide index= 2.2005 Wavelength= .633µm
Outer waveguide depth= 3.00 µm
TE modes



(a)



(b)

FIGURE [2,6] *Profiles for LUNEBURG OVERLAY LENSES*

(a) effective index profiles,
including 'classical' lens
(b) overlay film thickness profiles
for synthesis of (a)

As2S3 Film on Ti-diffused LiNbO3
y-cut, x-propagating

Film index= 2.6 Substrate index= 2.2028
Outer waveguide index= 2.2065 Wavelength= .633μm
Outer waveguide depth= 3.00 μm
TE modes

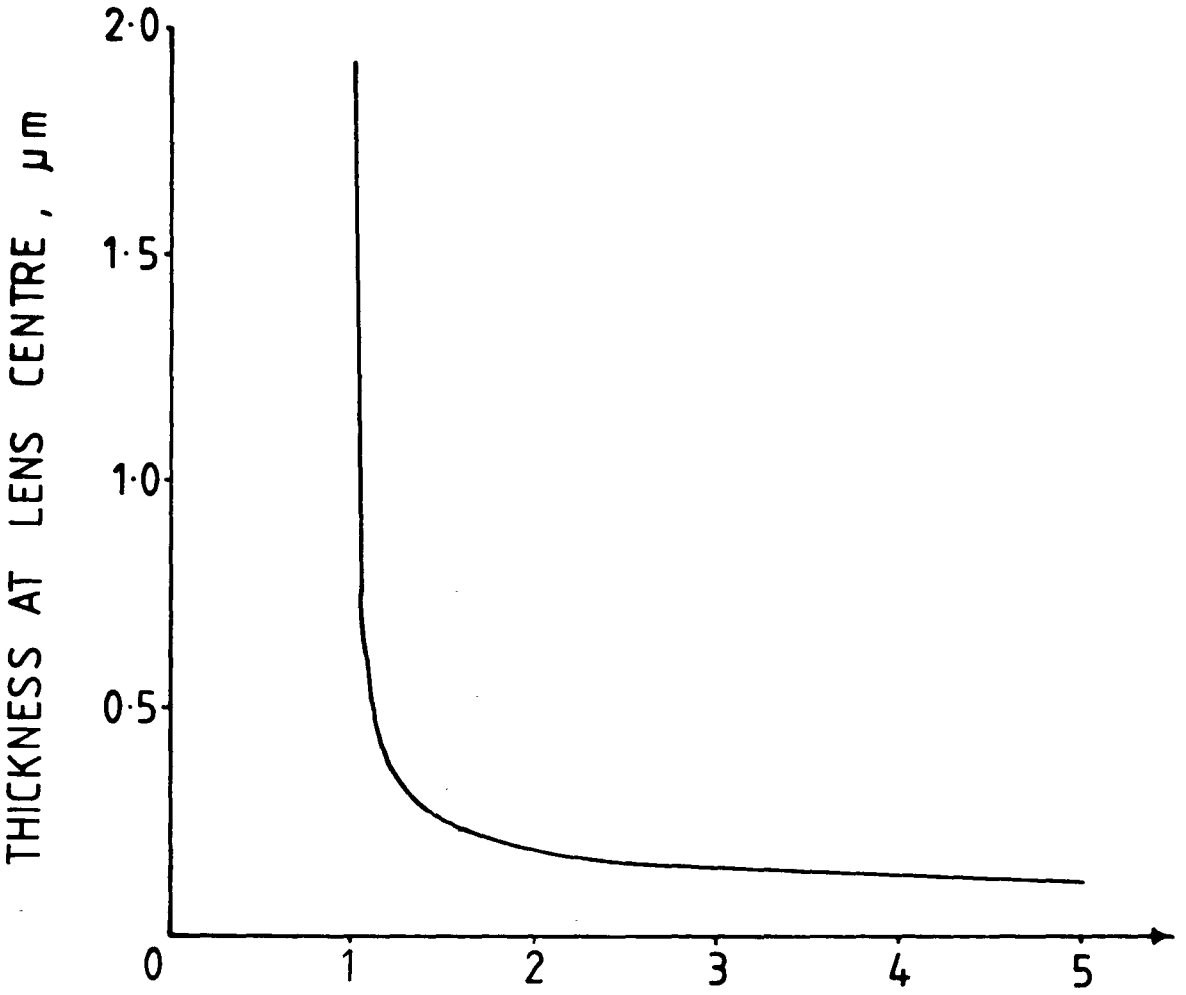


FIGURE [2,7] VARIATION OF THICKNESS AT CENTRE OF THIN-FILM LUNEBURG LENS WITH F-NUMBER (As_2S_3 on $Ti:LiNbO_3$)

and f/1.05 possess substantial regions of upward curvature. It will be shown in chapter three that such profile behaviour is impossible to obtain using conventional fabrication techniques.

The relationship between the film centre thickness and the required f-number may be changed substantially by choosing different combinations of materials for the substrate, ambient waveguide and overlayer. Figure [2.8] shows the effect of choosing As_2S_3 as the material for both the ambient waveguide and the lens overlayer, on a glass substrate. The minimum f-number that can be realised using such a combination increases, since the effective index difference that can be realized between the ambient waveguide and the overlayer is substantially reduced.

The lens profiles of Figures [2.5] and [2.6] were computed using the dispersion curve of the lowest order TE mode in Figure [2.4] . It can be seen from Figure [2.4] that overlay film thicknesses of greater than 0.5 wavelengths (approximately 315nm at a wavelength of 633nm) result in a substantially multi-moded regime. Figure [2.7] then indicates that only f-numbers greater than f/1.2 can be constructed for single-mode (or at most double-moded) operation.

In the light of the above considerations an f-number of f/2 seemed a reasonable design goal for the purposes of the present work.

2.5 Design of the aspheric geodesic lens

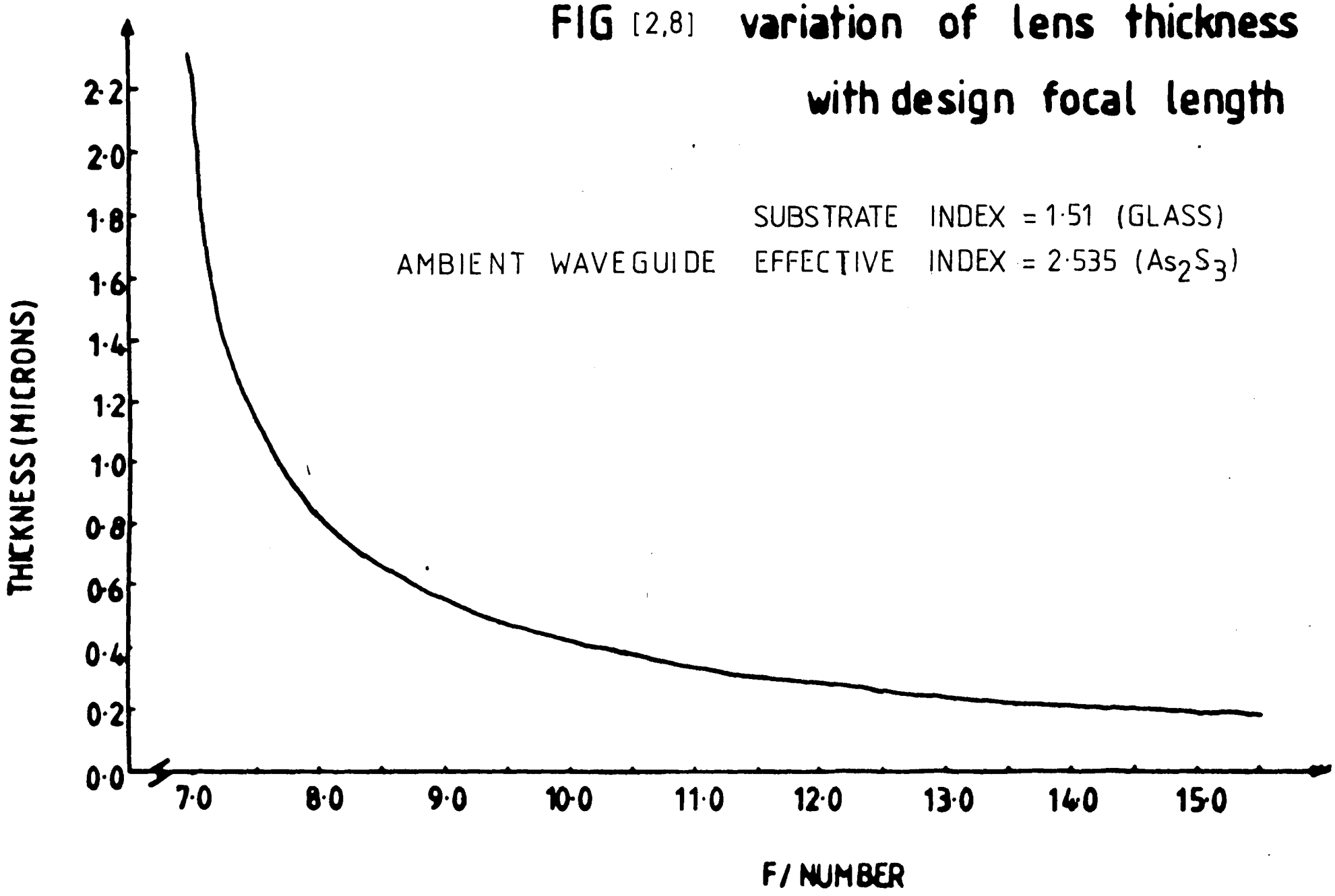
A solution to the problem of perfect focusing in a waveguide plane using a geodesic lens rather than an overlay lens has been given by Sottini et al [42]. Similar principles to the theory of Luneburg lenses apply, except that a geometrical rather than a refractive index profile is required. The design formulae will be given in this section without derivations.

Rays propagate along curved paths in geodesic lenses. The paths are the paths of minimum optical length between points on

FIG [2,8] **variation of lens thickness with design focal length**

SUBSTRATE INDEX = 1·51 (GLASS)

AMBIENT WAVEGUIDE EFFECTIVE INDEX = 2·535 (As_2S_3)



the rays, in accordance with Fermat's principle. Since the refractive index of the waveguide on the lens is assumed to be constant, the paths of minimum optical length coincide with the paths of minimum physical length, ie, the geodesics of the surface. Since the focusing properties of the lens are determined by the geometrical properties, the lens is achromatic, ie its properties do not change with the wavelength of the incident light. The geodesic lenses described by Sottini et al are similar to the Luneburg lenses previously in that they are capable of forming perfect geometrical images of two circles, one upon another. This is shown schematically in Figure [2.9] . Only imaging between circles external to the lens is considered here. The circle at radius a may be thought of as the object circle for convenience. The image circle is at radius b. In order for light to propagate smoothly from the ambient waveguide into the lens proper, a bridging region is incorporated between radial distances c and d, where c is the overall radius of the lens and the d is the radius of the inner focusing region.

For perfect focusing to occur, the meridional generating curve of the inner portion of the lens is given by:

$$z(r) = \int_0^r [1'(r)^2 - 1]^{1/2} dr \qquad (2.34)$$

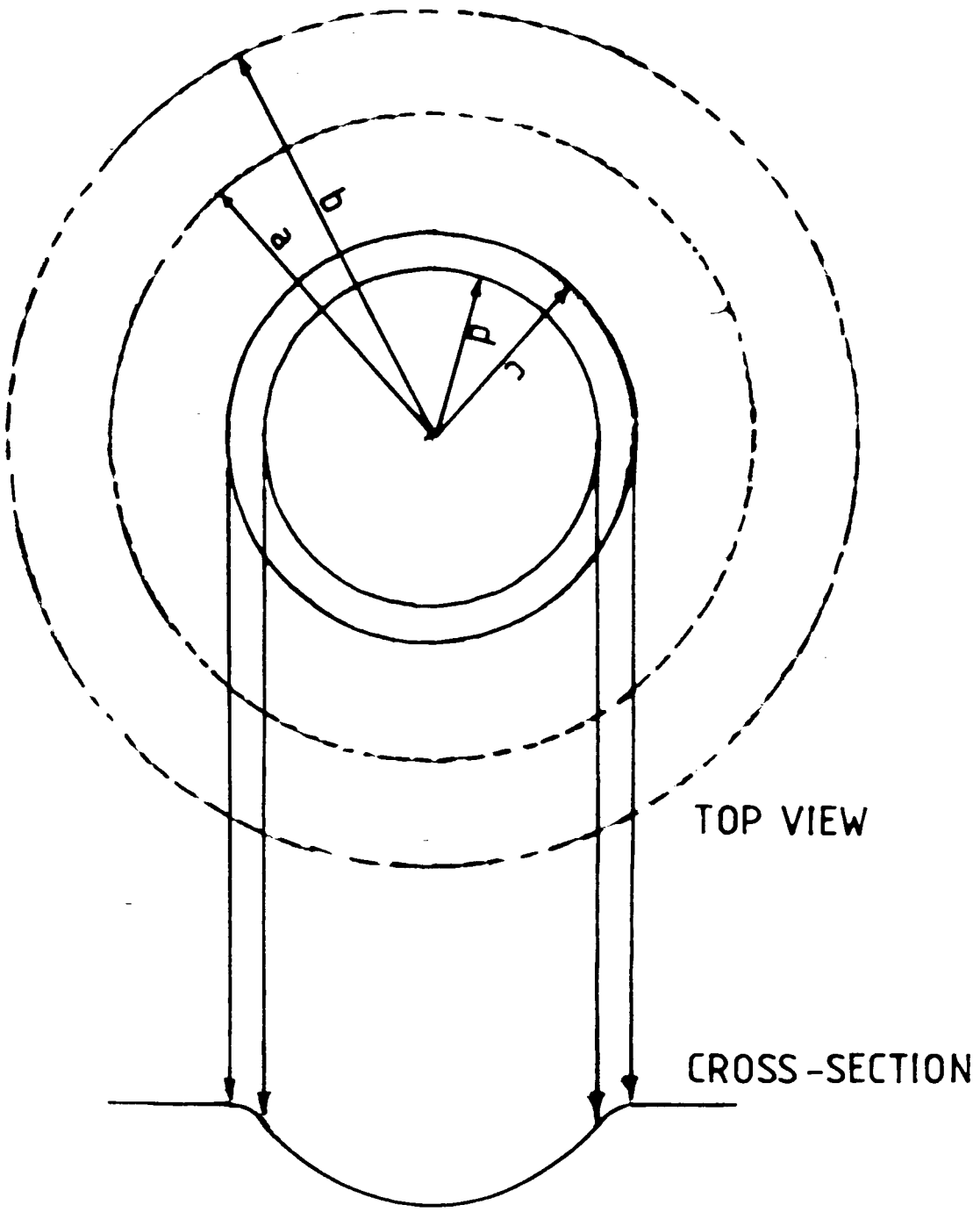
where 1'(r) is the first derivative of the arc-length, and may be shown to be given by:

$$1'(r) = -0.5\left(1.0 + \frac{c^2-r^2}{h^2}\right)$$

$$+ \frac{2(d^2-r^2)^{1/2}(c^2-d^2)^{1/2}}{\pi h^2}$$

$$- \frac{0.5}{\pi}\arcsin\left(\frac{r^2+b^2-2d^2}{b^2-r^2}\right)$$

$$- \frac{0.5}{\pi}\arcsin\left(\frac{r^2+a^2-2d^2}{a^2-r^2}\right)$$



FIGURE[2,9] ASPHERIC GEODESIC LENS FOR PERFECT IMAGING OF CONJUGATE CIRCLES

$$+ \frac{(r^2-c^2)}{\pi h^2} \arcsin\left(\frac{r^2+c^2-2d^2}{c^2-r^2}\right)$$

$$(0 \leqslant r \leqslant d) \qquad (2.35)$$

where h is defined by:

$$h = c\frac{(\pi - 2\gamma_{12} - \sin 2\gamma_{12})^{1/2}}{\gamma_{13} + \gamma_{14}} \qquad (2.36)$$

and:

$$\gamma_{12} = \arcsin(d/c)$$
$$\gamma_{13} = \arcsin(d/b)$$
$$\gamma_{14} = \arcsin(d/a)$$

The generating curve of the outer edge-rounding region is given by:

$$z(r) = z(d) + \int_d^r \left[\left(\frac{c^2-r^2}{h^2}\right)^2 + 2\left(\frac{c^2-r^2}{h^2}\right)\right]^{1/2} dr$$

$$(d < r \leqslant c) \qquad (2.37)$$

The focal length of this lens is given by the usual Gaussian formula:

$$f = \frac{ab}{a+b} \qquad (2.38)$$

and the maximum usable f-number is given by:

$$\text{f-number} = f/2d \qquad (2.39)$$

A feature of the lens profile is that the tangents of the generating curves of the inner and outer portion of the lens are matched at r=d, and the outer portion and the ambient waveguide are similarly matched at r=c. Continuity of the generating function is thus obtained. It will be seen in chapter seven of this work that the second derivative of the lens is not continuous and this can give rise to waveguide uniformity problems.

Doughty et al [43] fabricated lenses based on the above equations, with the intention of incorporating the lenses in an integrated optical spectrum analyzer (IOSA). The present author characterised the optical properties of these lenses from theoretical and experimental points of view. The parameters of the lens were:

a = 18.5mm
b = ∞
c = 5.0mm
d = 3.7mm

2.6 A simple geodesic design

The fabrication of the geodesic lenses of section 2.5 is a difficult task since the profiles are strongly aspheric. In some applications a simpler design based on spherical geometries can be utilised. Lenses based on spherical geometries generally suffer from image-defects or aberrations, but these can be minimised by operating the lenses at reduced apertures.

The form of the lens generating curve is shown in Figure [2.10]. As is the case with the aspheric geodesic lens, an edge-rounding region has to be incorporated in the design in order to effect a smooth low-loss transition between the lens proper and the ambient waveguide. A toroidal region is used for this purpose. The parameters R_s, a and Θ in Figure [2.10] completely

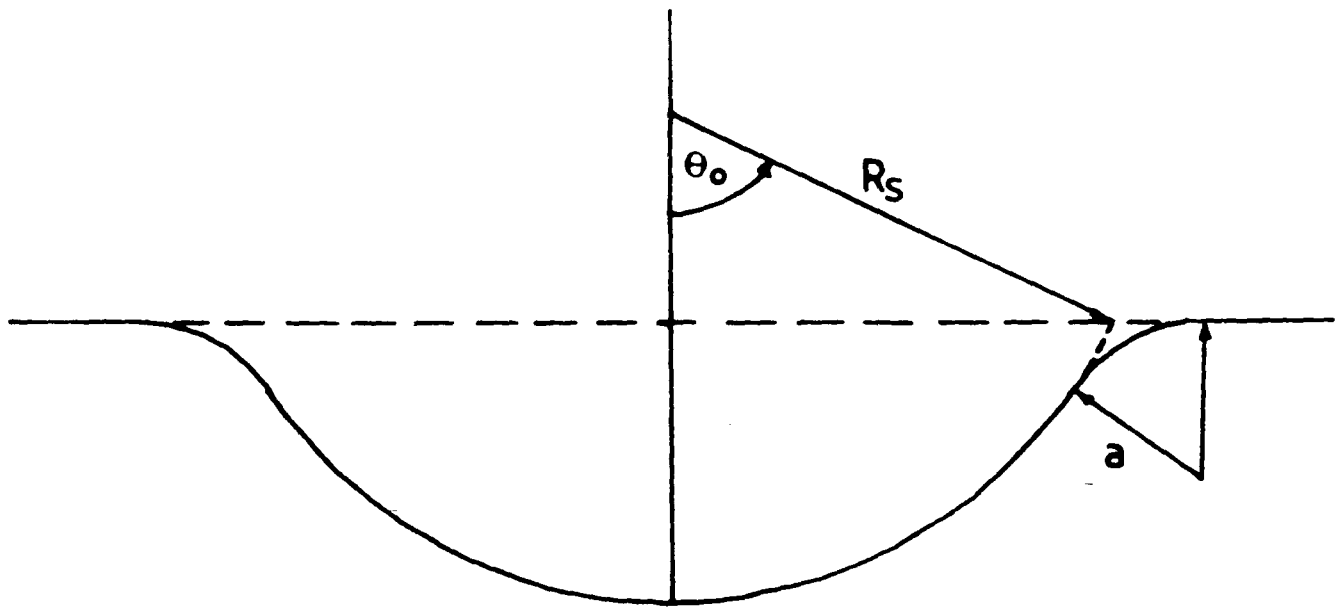


FIGURE [2,10] GEOMETRY OF SPHERICAL GEODESIC LENS WITH TOROIDAL EDGE-ROUNDING

define the lens. R_s is the radius of the generating circle for the inner portion of the lens, a is the radius of the edge-rounding region and Θ is the half-angle subtended by the inner portion of the lens. A lens of this design was fabricated by G.F. Doughty et al for GEC Marconi Research Centre, Chelmsford, England and was characterised by the author.

2.7 Conclusions

The design principles of inhomogeneous overlay Luneburg lenses, aspheric geodesic lenses and geodesic lenses based on spherical geometries have been given in this chapter. The constraints of single wavelength and single mode operation, focal length sensitivity to changes in overlay thickness, and profile realizability considerations restrict the feasible f-numbers obtained with As_2S_3/Ti:$LiNbO_3$ overlay lenses to a band between f/1.2 and f/4. An f-number of f/2 is considered to be a reasonable design goal. No such constraints apply to geodesic lenses and particular designs of both geodesic types have been specified which will be further considered in this thesis.

REFERENCES FOR CHAPTER TWO

[1] Born,M. and Wolf,E.
'Principles of Optics', Pergamon Press, Oxford, 1975

[2] Stavroudis,O.N.
'The Optics of Rays, Wavefronts and Caustics', Academic Press, New York, 1972

[3] Boivin,L.P.
Thin-film laser-to-fiber coupler, Appl. Opt., 13(2), 1974, 391-395

[4] Yu,Z.D.
Waveguide optical planar lenses in $LiNbO_3$-theory and experiments, Opt. Commun., 47(4), 1983, 248-250

[5] Righini,G.C., Russo,V., Sattini,S.(sic) and Toraldo di Francia,G.
Thin-film geodesic lens, Appl. Opt., 11(6), 1972, 1442-1443

[6] Skolnik,I.M.
'Radar Handbook', Ch. 10, McGraw-Hill, New York, 1970

[7] Anderson,D.B., Davis,R.L., Boyd,J.T. and August,R.R.
Comparison of optical-waveguide lens technologies, IEEE J. Quant. Electron., QE-13(4), 1977, 275-282

[8] Cornbleet,S.
Geometrical optics reviewed: a new light on an old subject, Proc. IEEE, 71(4), 1983, 471-503

[9] Luneburg,R.K.
'Mathematical Theory of Optics', Univ. of Calif. Press, Berkeley, California, 1964

[10] Kunz, K.S.
Propagation of microwaves between a parallel pair of doubly

curved conducting surfaces, J. Appl. Phys., 25(5), 1954, 642-653

[11] Toraldo di Francia,G.
A family of perfect configuration lenses of revolution, Opt. Acta., 1, 1955, 157-163

[12] Brown,J.
Microwave wide-angle scanner, Wireless Engineer, 1953, 250-255

[13] Gutman,A.S.
Modified Luneburg lens, J. Appl. Phys., 25(7), 1954, 855-859

[14] Morgan, S.
General solution of the Luneburg lens problem, J. Appl. Phys., 29(9), 1958, 1358-1368

[15] Miller,S.E.
Integrated optics: an introduction, Bell Sys. Tech. J., 48(7), 1969, 2059-2069

[16] Ulrich,R. and Martin,R.J.
Geometrical optics in thin-film light guides, Appl. Opt., 10, 1971, 2077-2085

[17] Van Duzer,T.
Lenses and graded films for focusing and guiding acoustic surface waves, Proc. IEEE, 58(8), 1970, 1230-1237

[18] Shubert,R. and Harris,J.H.
Optical guided-wave focusing and diffraction, J. Opt. Soc. Am., 61, 1971, 154-161

[19] Suematsu,Y., Furuya,K. and Kambayashi,T.
Focusing properties of thin-film lenslike light guide for integrated optics, Appl. Phys. Lett., 23(2), 1973, 78-79

[20] Zernike,F.
Luneburg lens for optical waveguide use, Opt. Commun., 12(4),

1974, 379-381

[21] Tien,P.K., Riva-Sanseverino,S., Martin,R.J. and Smolinsky,G.
Two-layered construction of integrated optical circuits and formation of thin-film prisms, lenses and reflectors, Appl. Phys. Lett., 24(11), 1974, 547-549

[22] Anderson,D.B., Boyd,J.T., Hamilton,M.C. and August,R.R.
An integrated-optical approach to the Fourier transform, IEEE J. Quant. Electron., QE-13(4), 1977, 268-275

[23] Southwell,W.H.
Index profiles for generalized Luneburg lenses and their use in planar optical waveguides, J. Opt. Soc. Am., 67(8), 1010-1014 (1977)

[24] Yao,S.K. and Anderson,D.B.
Shadow-sputtered diffraction-limited waveguide Luneburg lenses, Appl. Phys. Lett., 33(4), 1978, 307-309

[25] Yao,S.K., Anderson,D.B., August,R.R., Youmans,B.R. and Oania,C.M.
Guided-wave optical thin-film Luneburg lenses: fabrication technique and properties, Appl. Opt., 18(24), 1979, 4067-4079

[26] Bryan,D.A., Chubb,C.R., Powers,J,K,, Tomaschke,H.E., Reed,W.R. and Dalke,E.A.
Development of a tantalum pentoxide Luneburg lens, Proc. SPIE, 321, Integrated Optics II, 1982, 2-8

[27] Hatakoshi,G., Inoue,H., Naito,K., Umegaki,S. and Tanaka,S.
Optical waveguide lenses, Opt. Acta, 26(8), 1979, 961-968

[28] Busch,J.R., Wood,V.E., Kenan,R.P. and Verber,C.M.
Evaporated As_2S_3 Luneburg lenses for $LiNbO_3$:Ti optical waveguides, Optical Information Processing for Aerospace Applications, NASA Conference Publication 2207 (NASA Scientific and Technical Information Branch,1981), 251-261

[29] Colombini,E.
Design of thin-film Luneburg lenses for maximum focal length control, Appl. Opt., 20(20), 1981, 3589-3593

[30] Rinehart,R.F.
A solution of the problem of rapid scanning for radar antennae, J. Appl. Phys., 19, 1948, 860-862

[31] Rudduck,R.C. and Walter,C.H.
A general analysis of geodesic Luneburg lenses, IRE Trans. Ant. Prop., 1962, 444-450

[32] Righini,G.C., Russo,V., Sottini,S. and Toraldo di Francia,G.
Geodesic lenses for guided optical waves, Appl. Opt., 12(7), 1973, 1477-1481

[33] Spiller,E. and Harper,J.S.
High-resolution lenses for optical waveguides, Appl. Opt., 13(9), 1974, 2105-2108

[34] Vahey,D.W. and Wood,V.E.
Focal characteristics of spheroidal geodesic lenses for integrated optical processing, IEEE J. Quant. Electron., QE-13(4), 1977, 129-134

[35] Betts,G.E., Bradley,J.C., Marx,G.E., Schubert,D.C. and Trenchard,H.A.
Axially symmetric geodesic lenses, Appl. Opt., 17(15), 1978, 2346-2351

[36] Chen,B. and Ramer,O.G.
Diffraction-limited geodesic lens for integrated optic circuit, IEEE J. Quant. Electron., QE-15(9), 1979, 853-860

[37] Kassai,D. and Marom,E.
Aberration-corrected rounded-edge geodesic lenses, J. Opt. Soc. Am., 69(9), 1979, 1242-1248

[38] Bradley,J.C., Hutcheson,L.D., Kellner,A.L., Malarkey,E.C., Mergerian,D. and Pautenius,R.P.
Geodesic lens performance characteristics, Proc. SPIE, 239, Guided-wave Optical and Surface Acoustic Wave Devices, Systems and Applications, 1980, 84-89

[39] Mergerian,D., Bradley,J.C., Pautenius,R.P, Hutcheson,L.D., Kellner,A.L, Malarkey,E.C. and Marx,G.E.
Diamond-turned aspheric geodesic waveguide lenses in Lithium Niobate, Tech. Digest, Topical Meeting on Integrated and Guided-wave Optics, Nevada, USA, 1980, me4.1-4.4

[40] Myer,J.H. and Ramer,O.G.
Diffraction-limited geodesic lens: a search for substitute contours, Appl. Opt., 20(3), 1981, 412-416

[41] Lilienhof,H.J., Pantschew,B. and Schulz,D.
Printed geodesic lenses in glass with ion-exchanged film waveguides, 18(8), 1982, 344-345
ELECTRON. LETT.

[42] Sottini,S., Russo,V. and Righini,G.C.
General solution of the problem of perfect geodesic lenses for integrated optics, J. Opt. Soc. Am., 69(9), 1979, 1248-1254

[43] Doughty,G.F., DeLaRue,R.M., Singh,J., Smith,J.F. and Wright,S.
Fabrication techniques for geodesic lenses in Lithium Niobate, IEEE Trans. Comp. Hyb. Man. Tech., CHMT-5(2), 1982, 205-209

[44] Doughty,G.F., DeLaRue,R.M., Finlayson,N., Singh,J. and Smith,J.F.
Integrated optical microwave spectrum analyser (IOSA) using geodesic lenses, Proc. SPIE, 369 Max Born Conference, Edinburgh, Scotland, 1982, 705-710

[45] Kogelnik,H.
An introduction to integrated optics, IEEE Trans. Micro. Theory

Tech., MTT-23(1), 1975, 2-14

[46] Wong,S.W., private communication, 1984:

Ms. Wong and the author, using a multilayer model for calculating the optical propagation characteristics, calculated the values of TE mode effective refractive indices for three models of a Ti-diffused waveguide, each with an overlay film of As_2S_3. The models of the Ti-diffused waveguide were step, Gaussian and exponential with the 1/e depths of the latter two models both held equal to the depth of the step. The effective refractive indices of the composite waveguide systems for five values of overlay film thickness are shown in the Table below. Little variation is seen and the conclusion may be made, therefore, that the actual form of the Ti-diffused waveguide will have only a very small bearing on the lens profiles calculated in this chapter. The reason is, of course, that the refractive index difference between the overlay film and the outer waveguide is much greater than the refractive index difference between the outer waveguide and the substrate. If the differences were comparable in magnitude, such as might be found were the outer waveguide to be fabricated using the proton-exchange technique, for example (see chapter seven), then the actual form of the refractive index profile would certainly be important.

TABLE [2.4] Normalized effective indices for TE0 mode in As2S3/Ti:LiNbO3 structure for three ambient waveguide refractive index profiles

Overlayer thickness, um	AMBIENT REFRACTIVE INDEX PROFILE		
	STEP	GAUSSIAN	EXPONENTIAL
0.06328	1.000128	1.000193	1.000155
0.30058	1.128813	1.129252	1.129422
0.63280	1.163570	1.164026	1.164202
1.01248	1.172347	1.172804	1.172980
1.40798	1.175458	1.175916	1.176093

CHAPTER THREE

THE SHADOW-MASKING PROBLEM

CHAPTER THREE

THE SHADOW-MASKING PROBLEM

3.1 Introduction

The design of overlay Luneburg lenses was investigated in some detail in chapter two. In the present chapter necessary theory relevant to the problem of actually fabricating these lenses as closely as possible to a given design will be presented. All studies of overlay Luneburg lenses reported in the literature have thus far used a shadow-masking technique together with condensation-in-vacuum from a vapour-phase environment [1-4], but these have varied considerably in implementation ranging from the use of a simple circular mask of truncated cone cross-section to masks of aspheric geometry machined on NC lathes.

In some respects the problem is similar to geometrical optics in that the trajectories of molecules travelling in the fabrication chamber can be closely approximated by straight lines. Unfortunately, as in geometrical optics, extreme complications can set in due to the many degrees of freedom which a given system can possess. Essentially one desires to predict the paths of molecules impinging upon a planar substrate such that an apertured mechanical blocking mask may create a geometrical shadow which will allow the film growth profile to be closely controlled. It is therefore necessary to have an accurate description of the pattern of the molecular flux emanating from the source of material.

In section 3.2 of this chapter, the classical models of vacuum evaporation and deposition processes will be outlined. The possibility of approximating desired overlay lens profiles using the film thickness variation predicted by one of these models is investigated in section 3.3 . The approximations obtained are relatively poor. Yao's method of approximating the profiles in a sputtering environment by introducing mechanical masks between the source of deposited material and the target substrate is

discussed in section 3.4 . Yao's model is inappropriate for vacuum evaporation environments. A shadow-masking model appropriate to evaporation environments is presented in section 3.5, and the approximate lens profiles computed using the model are given in section 3.6 .

3.2 Holland and Steckelmacher models

Before embarking on a detailed study of the manipulation of molecular radiation patterns, it is worthwhile to return to a classic paper by Holland and Steckelmacher [5] which became the fundamental reference for most of the work on thin-film uniformity. Three important concepts may then be brought to light: (a) the nature of a point source; (b) the nature of a directed-surface-source; and (c) the significance of the equilibrium molecular mean free path. Only evaporated films are dealt with here, so that, for example, rf sputtering by heavy-ion bombardment is not necessarily described. A necessary condition to be satisfied is that the evaporation rate be sufficiently low so that the effect of molecular collisions in the vicinity of the vapour source may be neglected. One also assumes that the evaporation rate remains uniform throughout the deposition. A further assumption is that the temperature of the source is everywhere uniform, ie, the source is isothermal.

A point source will have even emission in all directions in space. The amount of material dM passing through a solid angle $d\omega$ per unit time is, therefore, given by:

$$\frac{dM}{d\omega} = \frac{m}{4\pi} \qquad (3.1)$$

where m is the total mass of material emitted from the source per unit time. Consequently, the amount of material condensing on a surface element dS of a receiving substrate which has its normal at an angle Θ to the direction of the source from the element will be given by

$$dM = \frac{m \cos\theta \, dS}{4\pi r^2} \qquad (3.2)$$

where r is the distance from the source to the surface element.

A directed-surface source is a small, planar emitting surface which emits to one side only. A directed-surface-source has an additional directionality factor, cos ϕ:

$$\frac{dM}{d\omega} = \frac{m \cos \phi}{\pi} \qquad (3.3)$$

A proportionality constant of $1/\pi$ now appears because the source is planar. Equation (3.3) is known as Knudsen's cosine law. The amount of material condensing on a target substrate element, using a directed-surface source, is:

$$dM = \frac{m \cos\theta \cos\phi \, dS}{\pi r^2} \qquad (3.4)$$

If the deposited film has a density ρ and a deposited thickness per unit time d, then:

$$dM = \rho \, d \, dS \qquad (3.5)$$

and the film thicknesses condensing on the infinitesimal surface area dS are given by:

$$d = \frac{m \cos\theta}{4 \rho \pi r^2} \qquad (3.6)$$

and:

$$d = \frac{m \cos\theta \cos\phi}{\rho \pi r^2} \qquad (3.7)$$

for the point and directed-sources respectively. It is easy to show that a point source situated at the centre of a sphere would coat the wall of the sphere uniformly, whilst the directed-source would have to be situated on the sphere wall itself to obtain the

same effect. If a flat plate held directly above the source is considered as shown in Figure [3.1] then the above expressions lead to the following normalised deposited thicknesses:

$$d' = d/d_o = \frac{1}{(1 + (r/h)^2)^{3/2}} \qquad (3.8)$$

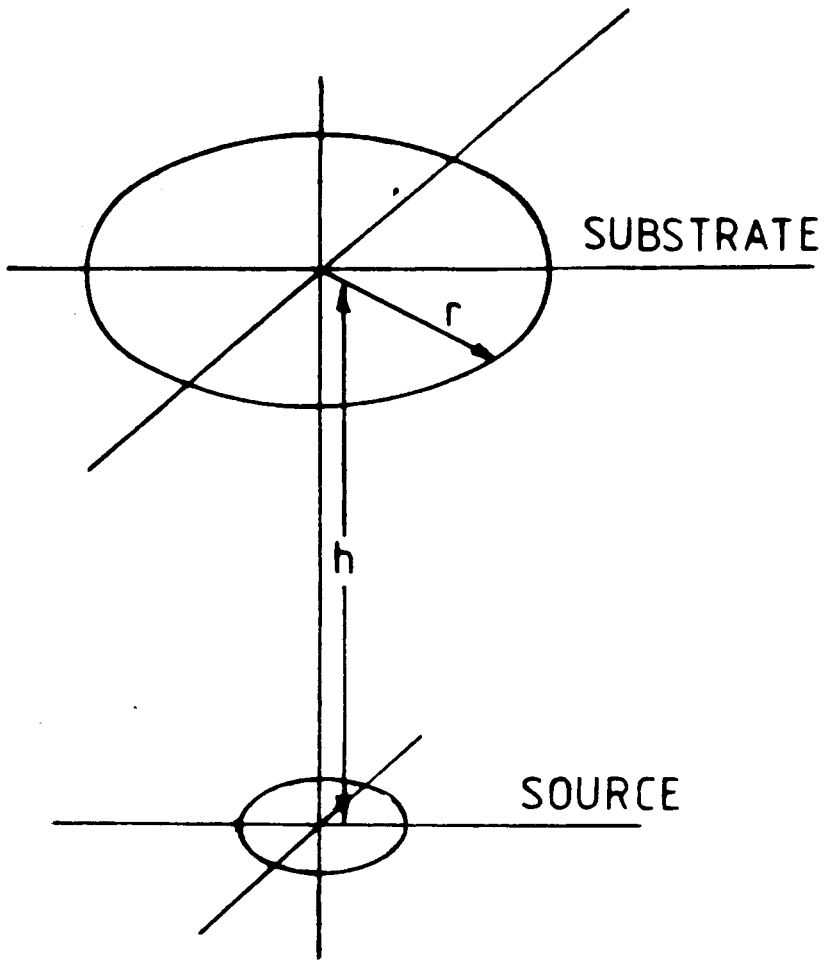
for the point source, and :

$$d' = \frac{1}{(1 + (r/h)^2)^2} \qquad (3.9)$$

for directed-surface source, where it is assumed that the flat plate is plane-parallel with the source. h is the perpendicular distance from source to substrate and r is the radial distance from the point on the substrate directly above the source where the deposited thickness is d_o. An eccentric rotation of one or both of the source and substrate, or a large source-to-substrate distance, is required to overcome the film non-uniformity implied by equations (3.8) and (3.9).

If the film condensation takes place in an evaporation system (such as a thermal evaporator or electron-beam evaporator) where the pressure is held below 10^{-4} torr, the molecular trajectories emanating from the source may be modelled as straight lines since the equilibrium mean free path of evaporated particles at such a pressure is generally much larger than both the source dimensions and the source-to-substrate distance. In sputtering environments pressures are higher and the equilibrium mean free path (and the non-equilibrium mean free path of sputtered molecules) is of the order of typical source-to-substrate distances (tens of mm) [6].

The Knudsen/Lambertian law is not always valid. It is appropriate for a small source, but real sources have finite and sometimes complicated geometry and are furthermore prone to secondary evaporation effects associated with adsorption of the evaporated particles to the source enclosure walls, followed by migration and re-evaporation. So called Clausing correction



FIGURE[3.1] SOURCE - SUBSTRATE GEOMETRY

factors have been introduced to quantify these effects [7], which usually cause increased 'beaming' or focusing of the molecular flux. A more rigorous presentation will be presented in section 3.5. Next, however, some simple calculations on Luneburg lens fabrication conditions based on the Knudsen directed-surface source will be presented.

3.3 Lens fabrication using a directed-surface source

In this section, the non-uniformity of film deposition implied by equation (3.9) will be investigated to see whether the overlay Luneburg lens thickness profiles calculated in chapter two could not be synthesised using such variation. The deviation between the desired and approximate profiles will be quantified. A mask situated on the substrate is assumed to truncate the approximate lens profile at the required full aperture.

Two lens designs are considered. One has an f-number of f/3, with a diameter of 4.5mm being somewhat arbitrarily selected so that the desired focal length is 13.5mm. The other lens has an f-number of f/9 and a diameter of 4.5mm, corresponding to a focal length of 40.5mm. The variable r, the radial distance in equation (3.9) is allowed to range between 0.0mm and 2.25mm at a constant value of h, the source-to-substrate distance. Curves of normalised deposited thickness (assuming a unity sticking coefficient) are shown in Figures [3.2] and [3.3] . Figure [3.2] shows that the source-to-substrate distance must be varied between 4.1mm and 6.3mm to obtain a 'reasonably close' fit to the f/3 lens, with 4.5mm to 9.0mm being the corresponding range for the f/9 lens as shown in Figure [3.3] These distances are obviously much smaller than those normally utilised in thin-film evaporation. If 'reasonably close' is defined as being determined by a least-squares deviation from the desired profile over the whole aperture, the best source-to-substrate distances are 4.3mm and 6.0mm respectively. Visually at least, the Knudsen curve approximations at these distances are not good, and other distances might seem more suitable. By inspecting Figure [3.3] more closely, for example,a strong case could be made for making

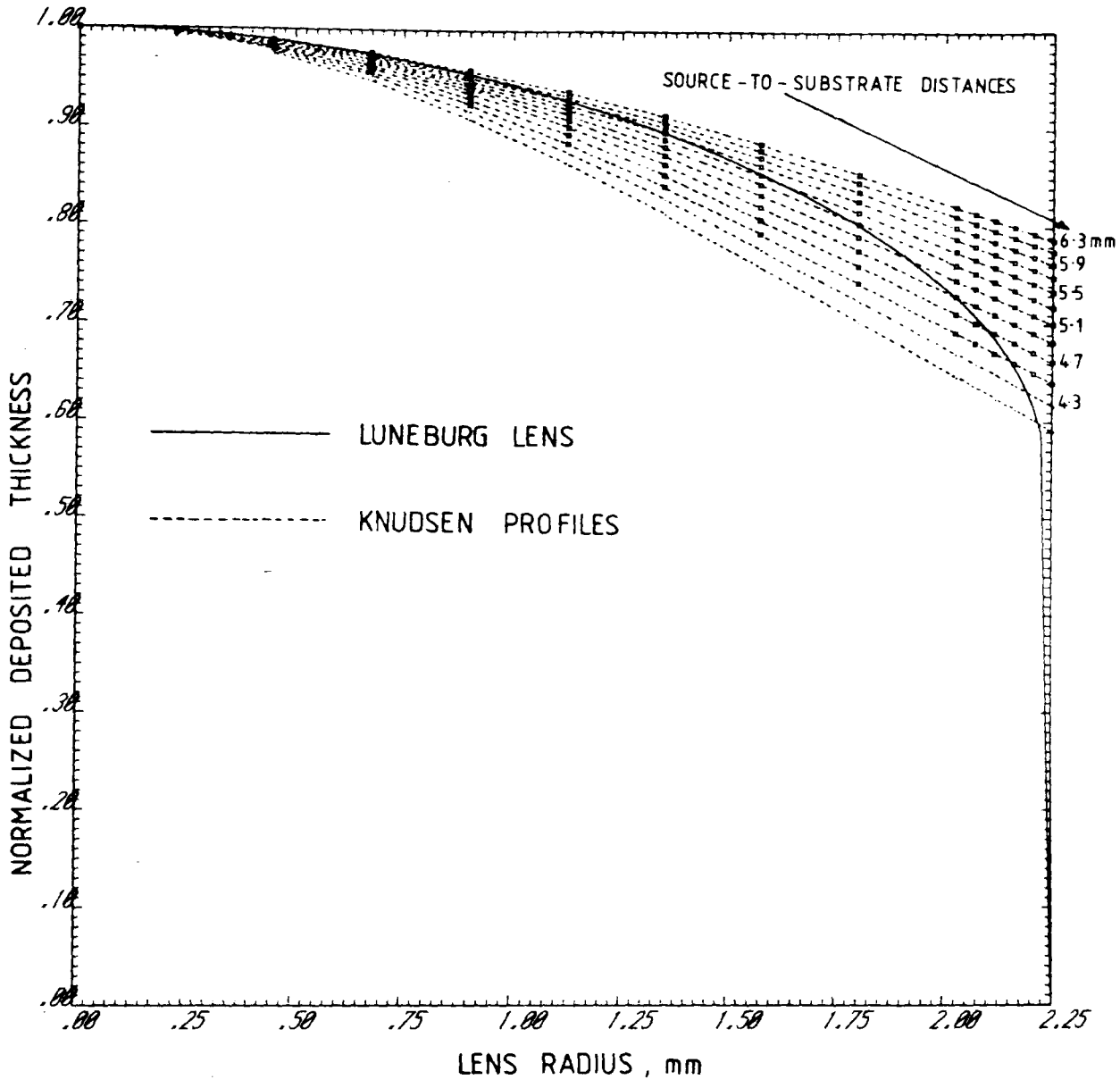


FIGURE [3.2] COMPARISON BETWEEN f/3 LUNEBURG LENS AND KNUDSEN LAW PROFILES

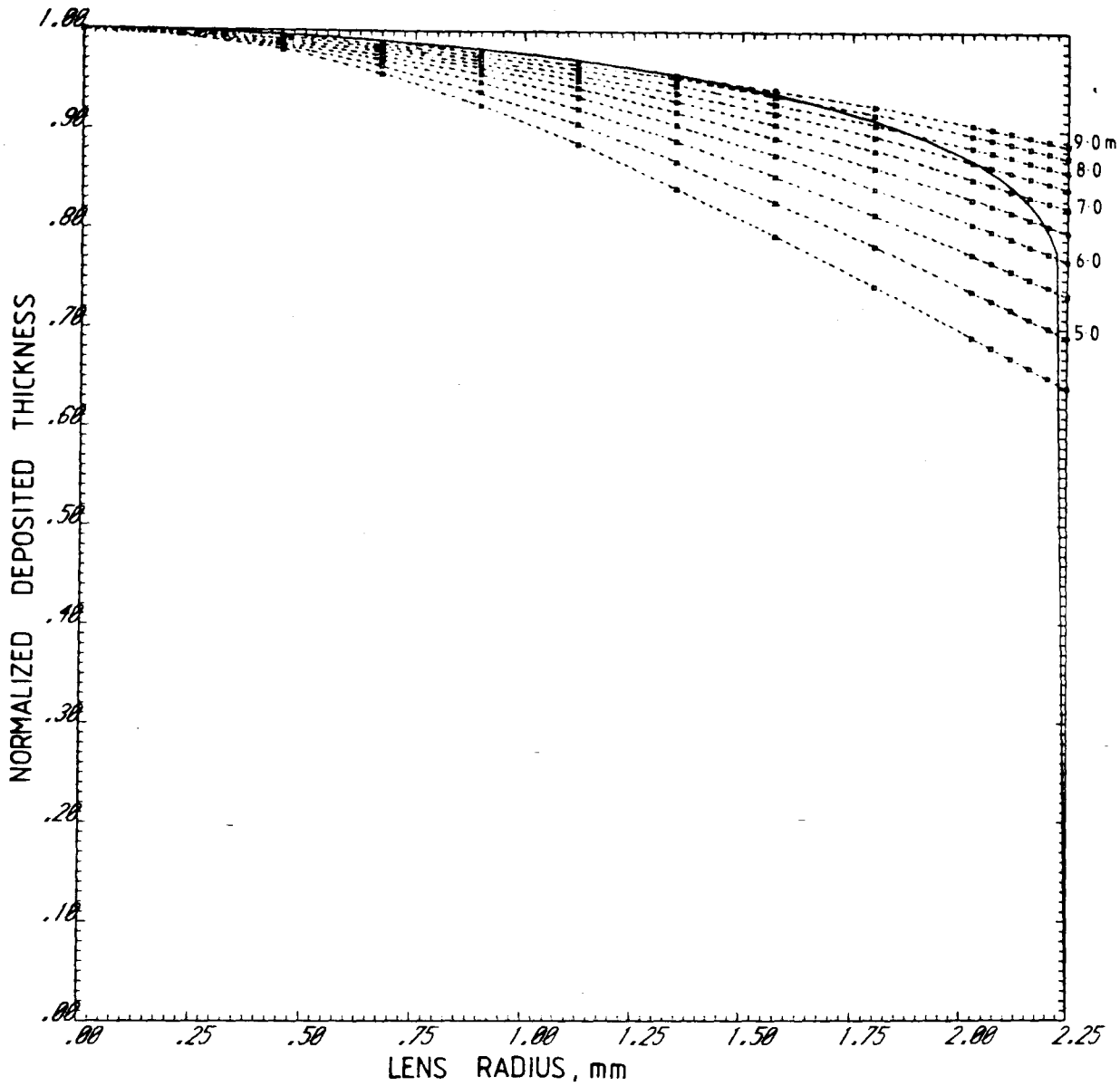


FIGURE [3,3] COMPARISON BETWEEN f/9 LUNEBURG LENS AND KNUDSEN LAW PROFILES

9.0mm the best source-to-substrate distance for fabricating the f/9 lens, since the profiles match closely up to a radius of 1.35mm. The least-squares criterion is, perhaps, disproportionately influenced by the points at the edge of the aperture.

Figure [3.4] shows a plot of whole-profile-deviation as a function of source-to-substrate distance for four lens designs given by f/1.5, f/3, f/6, f/9. The whole-profile-deviation is calculated by first squaring the deviation between the desired profile and the approximate profile at each point at which the desired profile is specified (29 in all) and then calculating the sum over all 29 points. Each design was investigated at four diameters d=4, 8, 12 and 16mm respectively. The minima of the whole-profile-deviation curves correspond to the optimum source-to-substrate distances for approximating the desired lens profiles. The graph is a complicated one, but several features are immediately apparent:

(i) the minimum whole-profile-deviation obtainable is approximately constant for a given f-number;

(ii) increasing the lens diameter, whilst holding the f-number constant has the effect of increasing the optimum source-to-substrate distance;

(iii) all the whole-profile-deviation minima fall in the source-to-substrate range 2-20mm, which are small distances, it may again be noted;

(iv) at large source-to-substrate distances the whole-profile-deviation is a constant for a given f-number. Lenses fabricated at these distances would be simple step-index lenses, analogous to spherical designs in bulk optics, and subject to image distortions (aberrations).

Conclusions to be drawn from the evidence presented so far are:

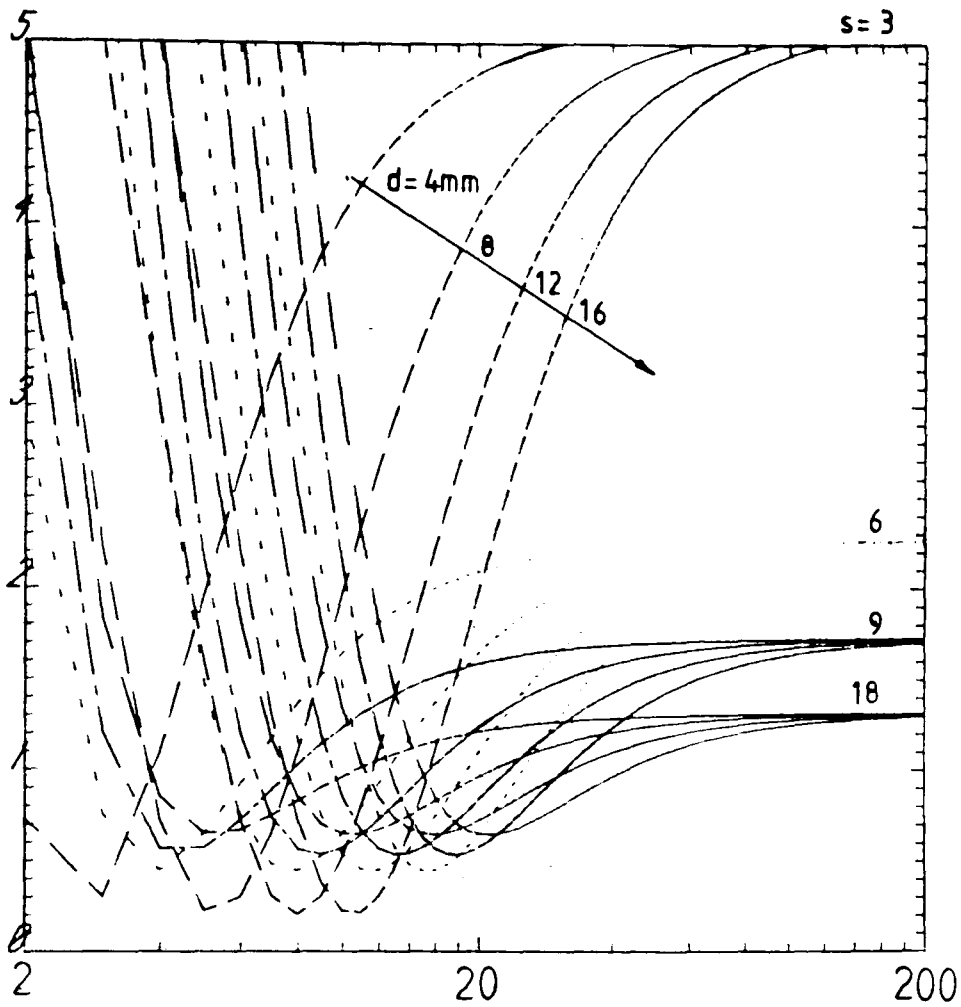


FIGURE [3.4] *FABRICATED LENS PROFILES*

error sum as function of source to substrate distance.

_ _ _ _ *S= 3.00*

. . . . . *S= 6.00*

_ . _ _ . _ *S= 9.00*

__ __ *S= 18.00*

f-number = s/2

each S-number investigated

over four diameters d=4, 8, 12, 16 mm

(a) the profile approximations obtained with a directed-surface-source of the type envisaged by Knudsen do not, in general, fit closely at all to the desired profiles. Substantial aberrations would be expected from any of the profiles investigated.

(b) Even if (a) were not the case, the source-to-substrate distances required are so small that it would be difficult to measure or control the deposited film thickness under experimental conditions.

(c) Furthermore, the small distances involved also result in a great sensitivity to substrate placement and alignment accuracy.

(d) Small distances also increase the possibility of material spitting from the evaporation source onto the substrate.

Evidently the deposited-film variation provided by a simple Knudsen directed-surface-source does not allow for easy or accurate fabrication of overlay Luneburg lenses.

Four other phenomena which can affect the deposited-film profiles may be either present in the evaporation process or introduced into it. These are:

(i) Source extension: An aggregate of directed sources is then available for modification of the observed molecular flux patterns.

(ii) The orifice geometry: A cylindrical orifice may be used to 'focus' the molecular beam.

(iii) Shadow masking: A mask could be interposed between the source and substrate to shape the profile. The shadow mask would not, in general, be situated on the substrate itself, although a second mask would be present on, or near, the surface

in order to define the diameter of the lens. The second mask could, furthermore, play a role in defining the deposition profile near the lens edge.

(iv) Secondary evaporation: These effects can considerably modify the molecular flux.

3.4 The Yao shadow-masking model

In his paper [8], Yao points out that the shadow-masking of thin films to achieve controlled variations in film thickness is often carried out on an ad hoc, trial-and-error basis and is dependent to a large extent on the skill and experience of the operator using the equipment. With the motivation of improving upon such a largely intuitive approach, Yao successfully formulated an algorithm for generating generalized mask profiles which would closely synthesize thin-film overlay Luneburg lens profiles. A computer-aided design method was used which had as design goals precision, predictability and reproducibility: all necessary attributes for lenses intended to be incorporated in practical optical processing systems.

After consideration of the generalized problem, ie one that makes allowances for

(a) variable source-to-substrate distances,

(b) finite, but complicated source geometries,

and (c) variable mask geometries,

Yao came to the conclusion that a computer model capable of covering all aspects of the general problem would be unwieldy and tedious. He restricted himself thereafter to the following cases:

(a) one where an extended source is used whose dimensions are large compared to the distance between source and

substrate,

(b) one where a shadow mask is included with a depth-to-diameter ratio so large that only small values of incident angle O need be considered.

Such conditions closely approximate those found in a sputtering environment, and a simple, efficient computing model is easily constructed. The conditions unfortunately do not correspond so closely to those found in a typical evaporation unit where the source is relatively small and secondary evaporation effects can occur. Neither does Yao invoke the possibility of beam-shaping by introducing an orifice. Secondary evaporation effects are fortunately negligible for materials such as As_2S_3 evaporated from the tantalum crucibles which were used in the author's work, due to the low binding energies between glasses and metals [9].

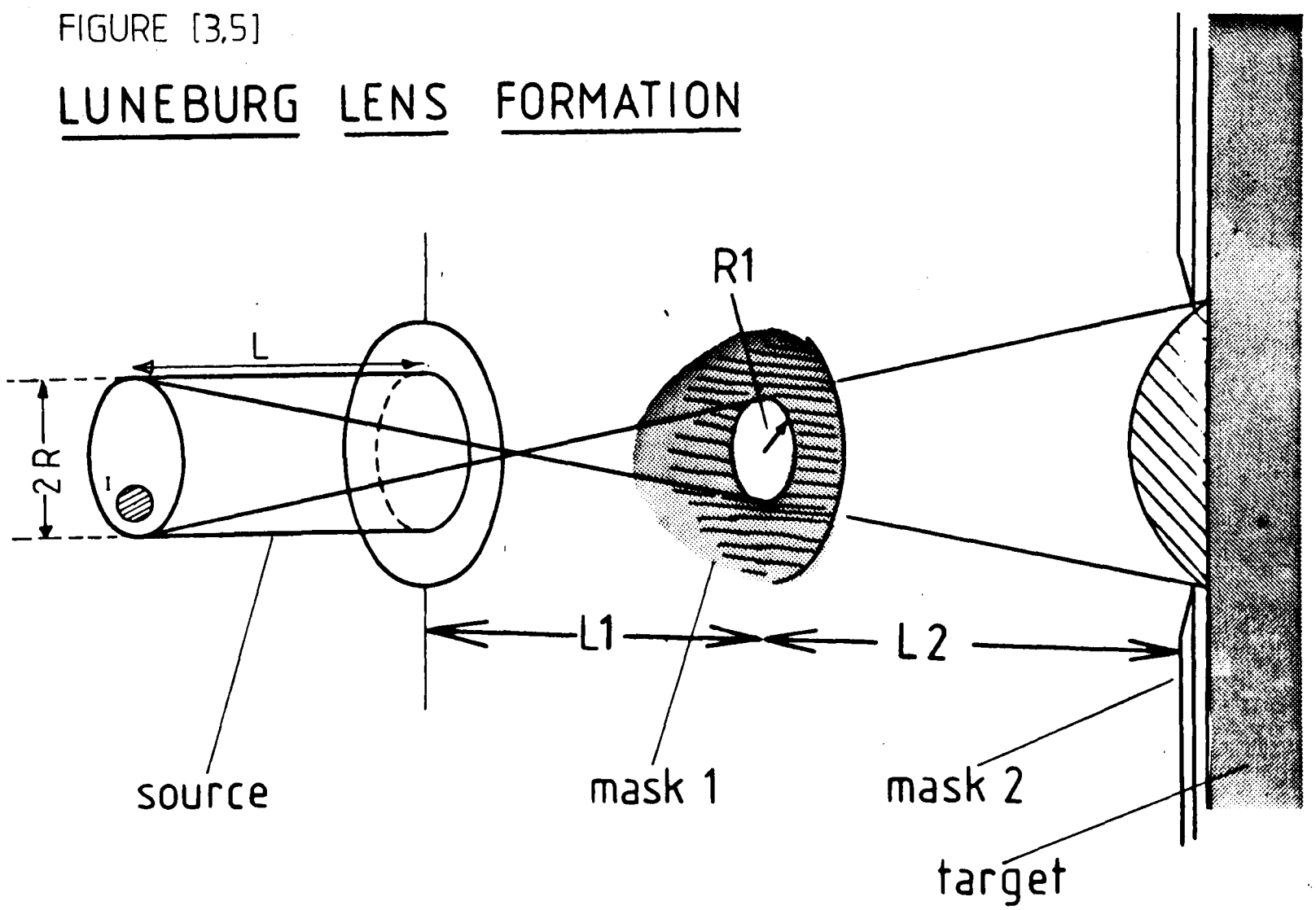
3.5 Masking the substrate

It has previously been noted that the film variation produced by a source on its own is insufficient for the purposes of fabricating lenses directly. A mask may be introduced to increase the available variation as illustrated in the schematic of Figure [3.5]. The mask is modelled as an infinitely thin sheet, completely opaque to the molecular flux except for a small disc of radius R1. Both source and mask aperture share the same axis of rotational symmetry, and the top surface of the source, the mask and the substrate are all plane-parallel. For mathematical convenience, the radius of the mask is assumed to be smaller than that of the source. A second mask is shown placed in close proximity to the substrate to improve the edge definition of the system.

The precise geometry of the situation is shown in Figure [3.6]. The source is a cylindrical orifice of radius R and length L. The shadow-mask has a radius R1 and is situated a distance L1 from the source. The target substrate is situated a distance L2 from the shadow-mask. The edge-definition mask is placed in close

FIGURE [3,5]

LUNEBURG LENS FORMATION



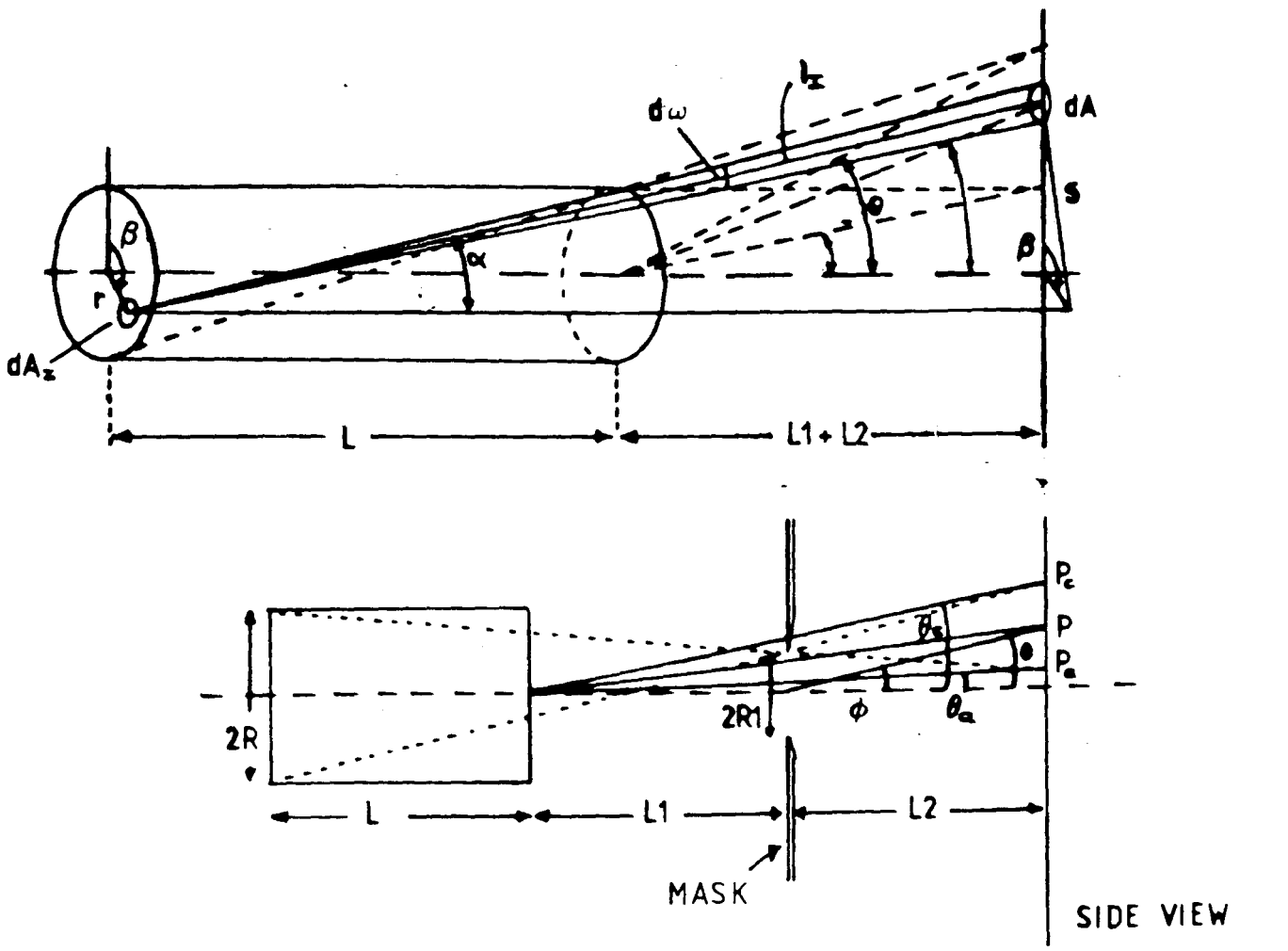


FIGURE [3.6] Shadow-masking geometry.
Cases (a), (b), (c) and (d) are referred to in the text.

proximity to the substrate and its influence may be neglected for the purposes of the following analysis. Each differential area dA of the orifice throat gives rise to a molecular contribution which follows the Knudsen law of equation (3.3). The total amount of material condensing at a point P on the target substrate characterized by the angle Θ is then given by:

$$J(\Theta) = k\int_A \cos\alpha \, d\omega \, dA \qquad (3.10)$$

where the angle α and the solid angle $d\omega$ are as shown on the Figure. It can be seen from the geometry of the Figure that:

$$d\omega = \frac{dA \cos\alpha}{l_I^2} \quad ,$$

$$s^2 = (L1+L2)^2\tan^2\Theta + r^2 - 2(L1+L2)\, r\, \tan\Theta \cos\beta \quad ,$$

$$\cos\alpha = \frac{(L + L1 + L2)}{l_I^2} \qquad \text{and}$$

$$l_I = \sqrt{\left[(L + L1 + L2)^2 + s^2\right]}.$$

Making the appropriate substitutions in equation (3.10):

$$J(\Theta) = \frac{k\ (L+L1+L2)^2}{[(L+L1+L2)^2+(L1+L2)^2\tan^2\Theta]^2} \times \int_A \frac{dA}{\left[1 + \frac{r^2 - 2(L1+L2)r\ \tan\Theta\cos\beta}{(L+L1+L2)^2 + (L1+L2)^2\tan^2\Theta}\right]^2} \qquad (3.11)$$

In most cases of interest, Θ is small and $r^2 \ll (L+L1+L2)^2$. The normalised deposition thickness at point P, characterised by the angle Θ, is then approximately given by:

$$J'(\Theta) \approx \frac{(L+L1+L2)^2}{[(L+L1+L2)^2+(L1+L2)^2\tan^2\Theta]^2} \quad A \qquad (3.12)$$

where $A = \int dA$ is the effective area of the orifice visible from point P.

Four cases may be identified as being of separate character, as shown in the figure. The geometrical shadow of the mask, as seen from point P, is cast onto the source, and the overlap defines the effective area.

(a) for points P characterised by angles less than Θ_a, where:

$$\Theta_a = \tan^{-1}(P_a/(L1+L2)) \qquad (3.13)$$

and:
$$P_a = \frac{R1\,(L+L1+L2)-RL2}{(L+L1)} \qquad (3.14)$$

the whole of the orifice area is seen.

(b) for points P characterised by angles greater than Θ_c, where:

$$\Theta_c = \tan^{-1}(P_c/(L1+L2)) \qquad (3.15)$$

and:
$$P_c = \frac{R1\,(L+L1+L2)+RL2}{(L+L1)} \qquad (3.16))$$

the orifice is entirely in shadow, and the deposited thickness is zero for all points in this region.

(c) for points P distinguished by angles Θ, where $\Theta_a < \Theta < \Theta_c$ the observable area will be defined by an integral to be evaluated below.

(d) as an adjunct to case (a), situations can arise where the shadow cast by the mask is smaller than the orifice area.

This occurs for long source-to-substrate distances and for small mask radii. In cases where the radius of the shadow is smaller than that of the source, points P at radial distances less than P_d, where:

$$P_d = \frac{(R-R1)(L+L1+L2)}{(L+L1)} - R \ , \qquad (3.17)$$

will see a reduced effective area.

The most general case is (b). The area A seen by the point on the target is given by

$$A = \int_{y'}^{R} (R^2-y^2)^{1/2} dy + \int_{y''}^{R'} (R'^2-y^2)^{1/2} dy \ , \qquad (3.18)$$

where the integration limits are as shown in Figure [3.6] . We note first that, from the geometry of the Figure:

$$\tan\phi = \frac{(L1+L2)\tan\Theta}{L2} \qquad (3.19)$$

$$= \frac{(y'-y'')}{(L+L1)} \qquad (3.20)$$

Also,

$$\frac{R'}{R1} = \frac{(L+L1+L2)}{L2} \qquad (3.21)$$

and $$R^2-y'^2 = R'^2-y''^2 \qquad (3.22)$$

Combining (3.22) with (3.20) yields the integration limits:

$$y' = \frac{(R'^2-R^2)}{2(L+L1)\tan\phi} - \frac{(L+L1)\tan\phi}{2} \qquad (3.23)$$

$$y'' = y' + (L+L1)\tan\phi \qquad (3.24)$$

Carrying out the integration of (3.18), the orifice area seen from point P becomes:

$$A = \frac{1}{2}\left\{ \frac{\pi}{2}(R'^2+R^2) + y'(R^2-y'^2)^{1/2} \right.$$
$$\pm R^2 \sin^{-1}(|\pm y'|/R)$$
$$\left. - y''(R'^2-y''^2)^{1/2} - R'^2 \sin^{-1}(y''/R') \right\} \qquad (3.25)$$

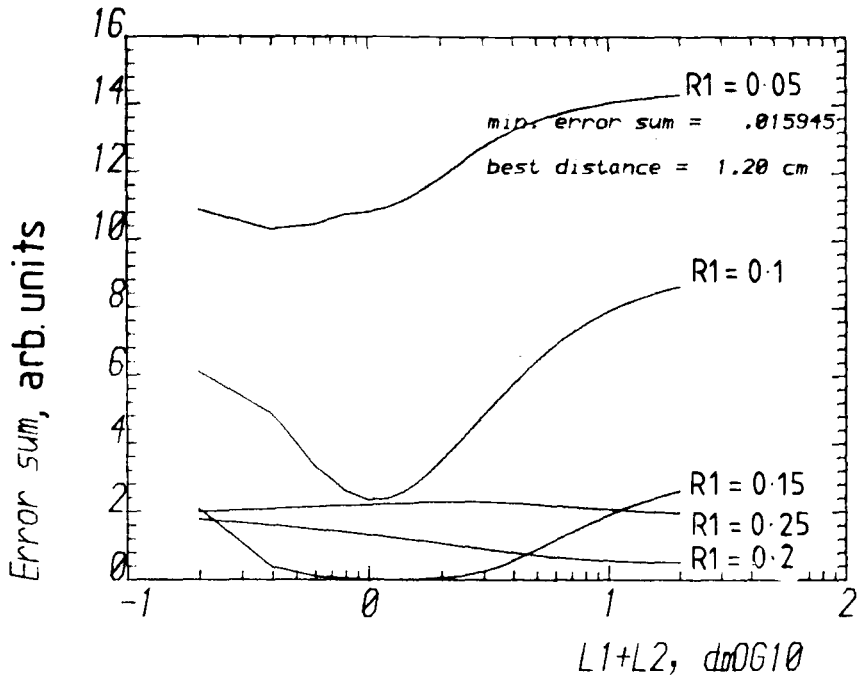
In case (a), $A = \pi R^2$.
In case (c), $A = 0$.
In case (d), $A = \pi R'^2$.

The overall distribution is affected therefore by the five variables L, R, L1, L2, and R1. Such a dependence allows modification of the fabrication conditions so that an approximation to the required lens profile is obtained.

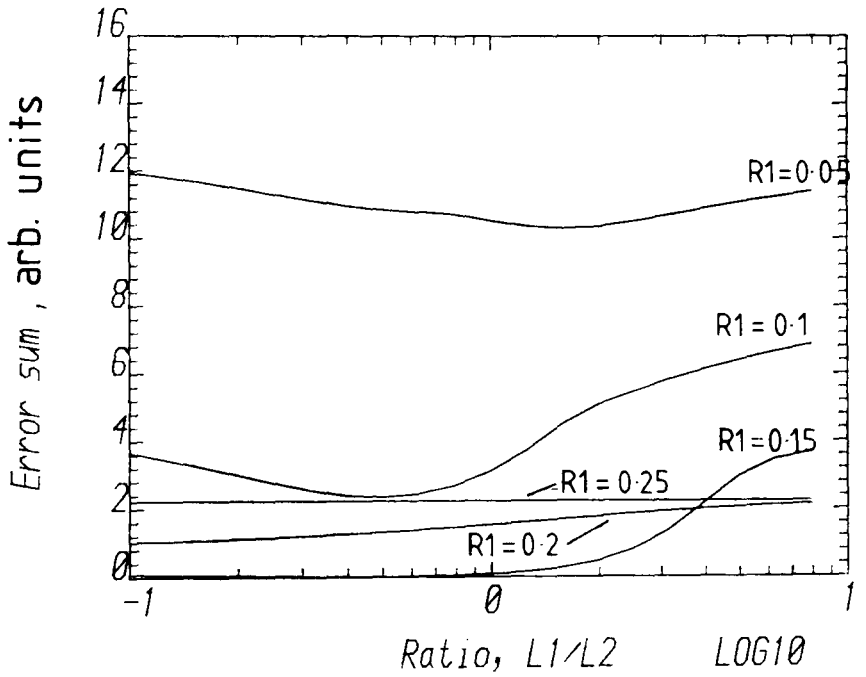
3.6 Lens approximations

A computer program has been written that calculates the whole-profile-deviation between the desired lens profile and the profile furnished by equation (3.12) over twenty seven profile points, as a function of the shadow-mask parameters. The source geometry is held constant while the mask radius R1, the sum (L1+L2) and the ratio (L1/L2) are allowed to vary over fixed ranges. The minimum whole-profile-deviation is then found and the values of L1, L2 and R1 which give rise to it are taken as the optimum values for lens fabrication. The source geometry is taken as fixed since only one type of source was used in the present study.

Figure [3.7] (a) shows the effects of varying the sum (L1+L2), ie the source-to-substrate distance, over a 2mm to 200mm range. An f/2 lens of diameter 4mm is being considered. The ratio (L1/L2) is held constant at the optimum value, and the error sum



(a)



(b)

Fig [3,7]

Plot of sum of squared errors
Luneburg lens, s-4, d=0.4cm
(a) as function of source-to-substrate distance
at optimum ratio = .398

(b) as function of ratio
at optimum sum = 1.200 cm

Error profiles measured at five mask radii:
R1 = 0.05, 0.1, 0.15, 0.2, 0.25 cm
Optimum = 0.15 cm

function is calculated at five mask radii ranging from 0.5mm to 2.5mm. In Figure [3.7] (b) the source-to-substrate distance is held constant, whilst the ratio of L1 to L2 (which determines the mask placement) is allowed to vary between .1 and 10.0 , again at five mask radii.

The curves are seen to be strongly dependent on the value of mask radius and a clear minimum is obtained for the radius 1.5mm. At larger mask radii than 1.5mm the error curves flatten considerably indicating the uniformity of deposition obtained with an extended, unmasked source. As the best mask radius is approached the width of the dips in the curves of Figure widen, indicating that the best obtainable profile is not a strong function of the source-to-substrate distance. Furthermore Figure [3.7] (b) indicates that the best profile (the lowest error sum) is not a strong function of mask placement either, with virtually any mask distance below 50% of the source-to-substrate distance being suitable.The mask radius is seen to be the key parameter.

Table [3.1] shows the optimum fabrication geometries for two f/2 lenses, one with a 4.0mm diameter and one with an 8.0mm diameter. As in the simple Knudsen case, the dimensions involved are not large, and poor thickness control and measurement would be expected in fabrication. A source of greater radius would undoubtedly improve matters, since more regions of the source would then be available for blocking by the shadow-mask. The mask and substrate could then be placed at greater distances from the source. This has been confirmed by other workers [2].

Figure [3.8] shows the desired thickness profiles of the two lenses and their expected approximations. The approximations , whilst not perfect, might be expected to produce reasonable results. The refractive index profile corresponding to the 4mm diameter lens, computed from the approximate thickness profile, is shown in Figure [3.9].

Lens diameter	Source- to- substrate distance	Source- to- mask distance	Mask radius
4 mm	12 mm	3.4 mm	1.5 mm
8 mm	16 mm	1.5 mm	2.25 mm

TABLE [3,1] Deposition geometries for s-4 lens

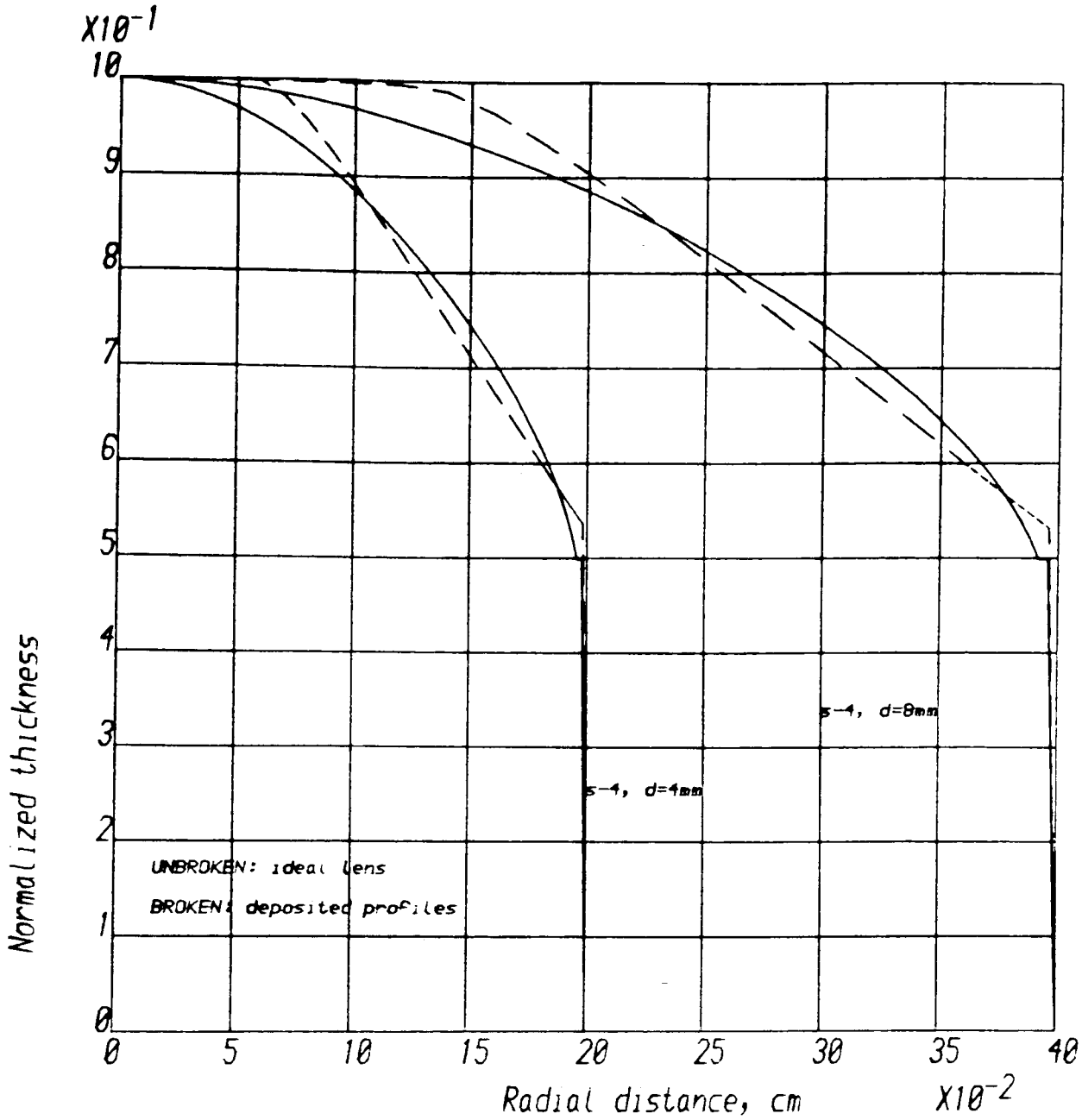


Fig [3,8] *Luneburg lens profiles*

and best obtainable approximations

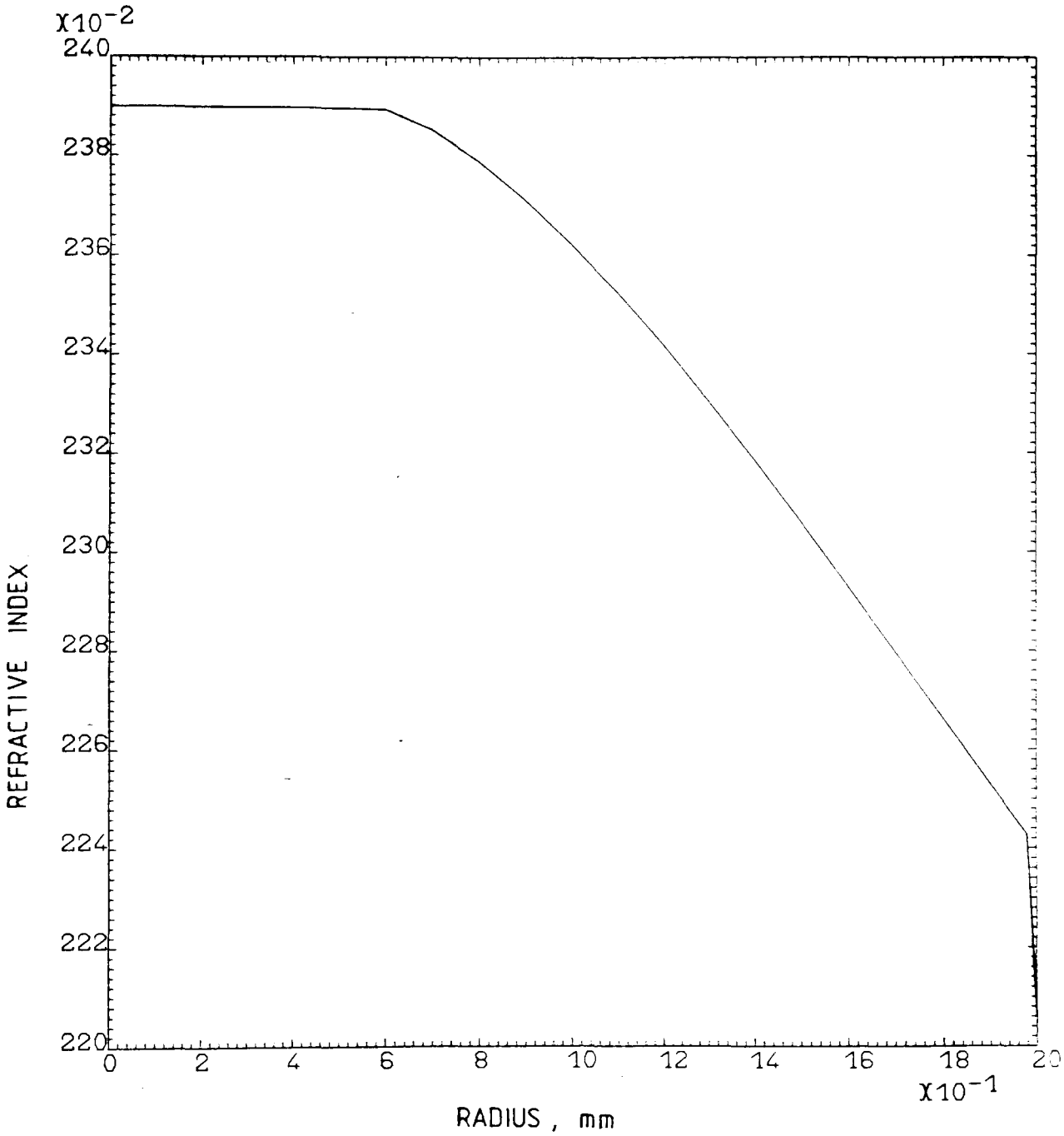


FIGURE [3,9] EFFECTIVE REFRACTIVE INDEX PROFILE OF LUNEBURG LENS APPROXIMATION

3.7 Conclusions

Vacuum evaporation models have been investigated. In order to obtain lens profiles close to the perfect Luneburg overlay profiles a theory of evaporation has been developed which allows for an extended, cylindrical source geometry and an infinitely thin shadow-mask of variable aperture, situated at variable distances from source and substrate. The theory predicts that reasonably good lens profiles can be obtained, if small source-to-mask and mask-to-substrate distances can be tolerated.

REFERENCES FOR CHAPTER THREE

[1] see chapter two, ref [25]

[2] " " " " [28]

[3] " " " " [27]

[4] " " " " [26]

[5] Holland,L. and Steckelmacher,W.
The distribution of thin films condensed on surfaces by the vacuum evaporation method, Vacuum, 11(4), 1953, 346-364

[6] Macleod,H.A.
"Thin-film Optical Filters", ch.9, Adam Hilger Ltd, London 1969

[7] Clausing,P.
Uber die Strumung sehr verdunnter Gase durch Rohren von beliebiger Lange, Ann. Phys., 12, 1932, 961-989

[8] Yao,S.K.
Theoretical model of thin-film deposition profile with shadow masking, J. Appl. Phys., 50(5), 1979, 3390-3395

[9] Hirth,J.P.and Moazed,K.L.
Nucleation processes in thin-film formation, Physics of Thin Films, 4, 1967, 97-136

CHAPTER FOUR

RAYTRACING

TRACING RAYS THROUGH INHOMOGENEOUS LENSES

4.1 Introduction

The geometrical optics of inhomogeneous media, non-planar media and thin-film waveguides enabled lens designs to be formulated in chapter two, some of which were expected to be perfect whilst others were known to be susceptible to imaging errors, ie aberrations. Practical lenses suffer from a variety of fabrication errors which include non-conformity to design profiles, misalignment errors, non-uniform film densities leading to effective index errors.

In such cases it is important to be able to model the passage of the light energy (and, to some extent, the passage of spatial information) through the particular optical system under investigation. The most widely used method remains that of tracing rays, usually in bundles, through lenses and observing the imaged results. As Stavroudis points out [1], an important reason for raytracing is that the presence of errors in the traced image

'...contributes distinctive geometrical characteristics to the structure of the image the appearances of which agree rather closely with what is seen in the laboratory.'

Miyamoto [2] provides impressive confirmation of the representational capabilities of ray plots, showing spot diagrams obtained by raytracing through bulk lenses side-by-side with photographs of images suffering from characteristic types of aberration.

Classically, the tracing of rays was a tedious business requiring repeated and laborious calculations. Optical designers

consequently developed a sophisticated aberration theory that extracted as much information as possible from the tracing of only a few rays. Modern digital computers have enabled many-ray systems to be computed and this has led to different methods of analysing the traced data. Nevertheless classical representations are still useful, particularly from a conceptual point of view.

Rays in isotropic systems are everywhere normal to the propagating wave phase-front [3]. The phase-fronts are parallel in the sense that subsequent surfaces are generated by stepping off equal distances along the ray normal to the original phase-front. As such the rays represent the curves along which the light energy is transported, within the limiting approximation of geometrical optics ie that of a vanishingly small wavelength. Later in this work it will be found that the energy contours derived from a more comprehensive theory (which deals with the field directly) do in fact closely resemble the rays seen in the present section.

In this chapter the fundamental equations of geometrical optics will be given. Two algorithms for tracing rays through inhomogeneous lenses will then be presented. One algorithm is used to trace rays through lenses of the inhomogeneous overlay type of lens, and the other is used to trace rays through the geodesic type. Results obtained using such ray-tracing algorithms will be presented which illustrate the behaviour of different lenses.

4.2 Pre-tracing

The requirements that lens ray-tracing procedures must meet are very simply stated. Given a single ray travelling in a prescribed direction which meets a lens at a prescribed intersection point , an algorithm is required which will obtain the corresponding position and direction at the point where the ray leaves the lens. Data collected for a bundle, or manifold, of rays provide a wealth of information on the imaging properties of a lens.

Ray-tracing concerns itself with establishing relationships between input and output quantities pertaining to the rays for the optical system through which the rays pass. A computational algorithm for tracing rays must be flexible enough to accomodate considerable variation of the input quantities. It is certainly desirable to control the following properties;

(a) the position in the object space of the source of rays (a point source is generally assumed, which in the case of a point situated at infinity gives rise to a parallel manifold of rays),

(b) the direction of each ray in the manifold with respect to the chosen axes,

and (c) the position and lateral extent of any stops in the system.

In addition, one might wish to include the capability of simultaneously tracing the rays at different wavelengths, to investigate the chromatic characteristics. In integrated optical applications the investigation of multi-mode effects could be important.

4.3 Fundamental equations of geometrical optics

The basic equation of geometrical optics is the 'eikonal' equation [3], eikonal being derived from the Greek word 'eikon', meaning image. The equation has the form:

$$(\mathrm{grad}(\Psi))^2 = n^2 \qquad (4.1)$$

in isotropic media, where $\Psi(\underline{r})$ is a real scalar function of position describing the optical path along a ray from a fixed point on the ray, and $n(\underline{r})$ is a real scalar function describing the refractive index distribution. The eikonal equation may be

derived by applying the limit of very small wavelength to the time-harmonic Maxwell electromagnetic equations for a non-conducting, isotropic medium free of currents and charges [3]. More familiarly, it may also be derived from the calculus of variations by starting with Fermat's principle. Such an approach formed the basis of Hamilton's work on the optical characteristic functions [4].

The surfaces given by:

$$\Psi(\underline{r}) = \text{constant} \qquad (4.2)$$

represent the geometrical phase-fronts. The rays are everywhere orthogonal to the phase-fronts in an isotropic medium and may be shown to point in the direction of the time-average Poynting vector. Consequently, the rays may be interpreted as being the curves along which energy is transported.

A unit vector $\underline{t}$ may be defined:

$$\underline{t} = \frac{\text{grad}(\Psi)}{n} \qquad (4.3)$$

which indicates the ray direction. If $\underline{r}(s)$ denotes the position vector of a point P on a ray, with s being the physical path length along the ray from a fixed position on it, then:

$$\underline{t} = \frac{d\underline{r}}{ds} \qquad (4.4)$$

from vector calculus [5], so that $\underline{t}$ is everywhere tangential to the ray. The eikonal equation may then be rearranged:

$$n \frac{d\underline{r}}{ds} = \text{grad}(\Psi) \qquad (4.5)$$

A simple physical interpretation of equation (4.5) is available. From vector calculus:

$$\frac{d\Psi}{ds} = \frac{\partial\Psi}{\partial x}\frac{dx}{ds} + \frac{\partial\Psi}{\partial y}\frac{dy}{ds} + \frac{\partial\Psi}{\partial z}\frac{dz}{ds}$$

$$= \left(\underline{i}\frac{\partial\Psi}{\partial x} + \underline{j}\frac{\partial\Psi}{\partial y} + \underline{k}\frac{\partial\Psi}{\partial z}\right)\cdot\left(\underline{i}\frac{dx}{ds} + \underline{j}\frac{dy}{ds} + \underline{k}\frac{dz}{ds}\right)$$

$$= \text{grad}\,\Psi \,.\, \frac{d\underline{r}}{ds}$$

$$= n \qquad (4.6)$$

Hence the well known result that the rate of change of the optical path with respect to the physical path along a ray is equal to the refractive index.

The rays are specified by means of the function Ψ in the eikonal equation. The rays can however be expressed directly in terms of the refractive index function by differentiating both sides of the eikonal equation (4.1) with respect to s. The equation:

$$\frac{d}{ds}\left(n\,\frac{d\underline{r}}{ds}\right) = \text{grad}\,(n) \qquad (4.7)$$

is obtained. Equation (4.7) is known as the differential equation of light rays or, more simply, the ray equation. It is the equation describing the paths followed by light rays as they traverse a material with refractive index n.

An important example is a medium with spherical symmetry, where the refractive index is a function only of the distance r from a fixed point O. In such a case it is possible to show that all the rays are plane curves situated in a plane through the origin. Along each ray:

$$n r \sin\phi = \text{constant} \tag{4.8}$$

where ϕ is the angle between the position vector $\underline{r}$ and the tangent vector $\underline{t}$ as shown in Figure [4.1]. Since $r\sin\phi$ represents the perpendicular distance d from the origin to the tangent, eq (4.8) may be written as:

$$nd = \text{constant} \tag{4.9}$$

Equation (4.9) is known as the formula of Bouguer.

An important relation concerning the ray curvature vector is often required in raytracing. The curvature vector $\underline{K}$ of a curve whose points are specified by $\underline{r}(s)$ is defined by:

$$\underline{K} = \frac{d^2\underline{r}}{ds^2}$$

$$= \frac{d\underline{t}}{ds} \tag{4.10}$$

Using the vector calculus result:

$$\frac{d(\underline{t}.\underline{t})}{ds} = 2\underline{t}.\frac{d\underline{t}}{ds} \tag{4.11}$$

which equals zero, since $\underline{t}$ has a constant magnitude, the vectors $\underline{K}$ and $\underline{t}$ are seen to be perpendicular to each other, with $\underline{K}$ pointing in the direction of the inward normal to the curve followed by the ray. Physically, since $\underline{t}$ has constant (unit) magnitude, $\underline{K}$ measures the rate of change of direction of $\underline{t}$. The magnitude of $\underline{K}$ is the reciprocal of the radius of curvature at the point $\underline{r}$.

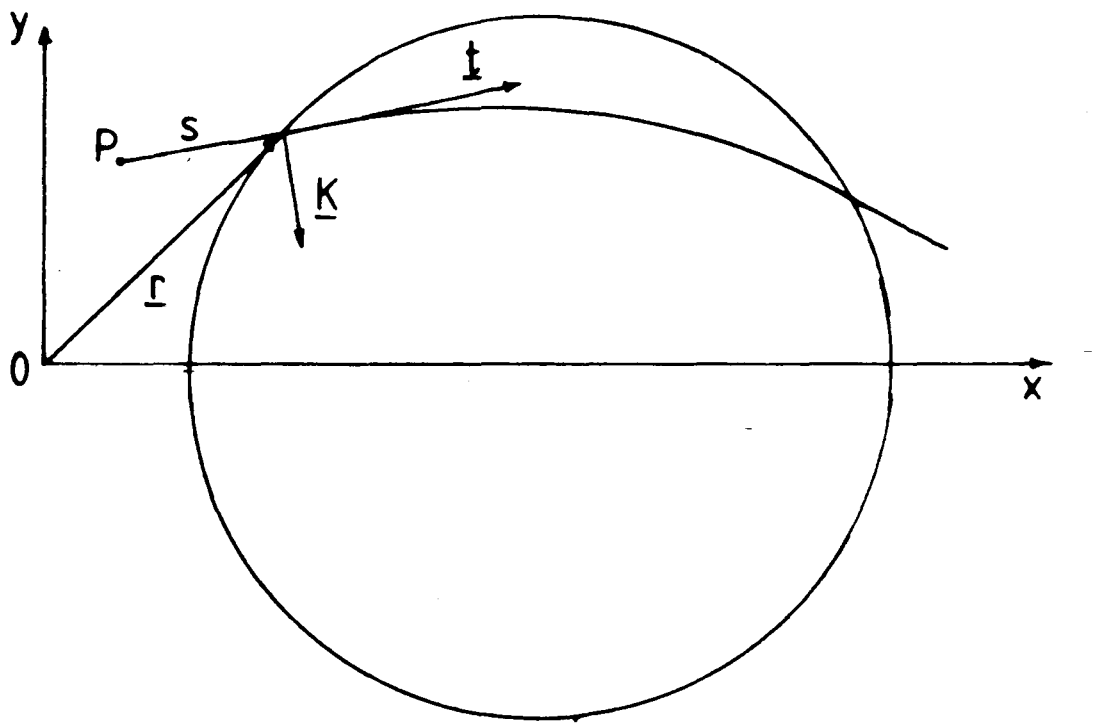
Carrying out the differentiation with respect to s in the ray equation (4.7) by means of the product rule:

$\underline{r}$ ray position vector

$\underline{t}$ ray direction vector

$\underline{K}$ curvature vector (measures rate of change of the direction of t)

s scalar distance along ray



FIGURE[4,1] GEOMETRY FOR TRACING RAYS THROUGH PLANAR INHOMOGENEOUS LENSES WITH CIRCULAR SYMMETRY

$$n \frac{d^2\underline{r}}{ds^2} + \frac{dn}{ds} \frac{d\underline{r}}{ds} = \text{grad}(n) \qquad (4.12)$$

Substituting $\underline{K}$ from (4.10) and rearranging:

$$n\underline{K} = \text{grad}(n) - \frac{dn}{ds} \underline{t}$$

$$= \text{grad}(n) - (\text{grad}(n).\underline{t})\ \underline{t} \qquad (4.13)$$

where the general relationship:

$$\frac{dg}{ds} = \text{grad}(g).\underline{t} \qquad (4.14)$$

has been used.

In a medium with spherical symmetry it is a simple matter to show that:

$$\text{grad}(n) = \frac{1}{r} \frac{dn}{dr} \underline{r} \qquad (4.15)$$

so that (4.13) becomes:

$$\underline{K} = \frac{1}{nr} \frac{dn}{dr} \left(\underline{r} - (\underline{r}.\underline{t})\ \underline{t} \right) \qquad (4.16)$$

The curvature vector is then fully specified in terms of the refractive index function.

Finally, the rate of change of $\underline{K}$ with respect to s is sometimes required in raytracing. This vector is defined by:

$$\frac{d\underline{K}}{ds} = \frac{d^3\underline{r}}{ds^3} \qquad (4.17)$$

In a spherically symmetric medium, $d\underline{K}/ds$ can be found explicitly

by differentiating both sides of equation (4.16):

$$\frac{d\underline{K}}{ds} = \frac{1}{r}\frac{df}{dr}\,(\underline{r}.\underline{t})\,[\underline{r} - (\underline{r}.\underline{t})\,\underline{t}]$$

$$- f\,[(\underline{r}.\underline{K})\,\underline{t} + (\underline{r}.\underline{t})\,\underline{K}] \qquad (4.18)$$

where:

$$f = \frac{1}{nr}\frac{dn}{dr} \qquad (4.19)$$

and:
$$\frac{df}{dr} = \frac{1}{nr}\left\{\frac{d^2n}{dr^2} - \frac{dn}{dr}\left(\frac{1}{r} + \frac{1}{n}\frac{dn}{dr}\right)\right\} \qquad (4.20)$$

The above equations enable rays to be traced through any spherically (or circularly) symmetric medium with a continuously differentiable refractive index function.

4.4 Tracing rays through inhomogeneous overlay lenses

4.4.1 Introduction

The trajectories followed by rays in inhomogeneous or non-uniform media have long been of interest. In certain cases where the refractive index function and its derivatives are restricted to special forms, exact analytical solutions are known [4] which usually depend on invoking symmetry considerations. The symmetries are generally of the cylindrical or spherical kind. Unfortunately, as pointed out by Moore [6], the assumed gradients of the index do not always resemble those of practical concern and the resulting mathematical solutions are not of great value in the optimisation of designs.

More useful methods involve approximations of the ray path and the refractive index gradient. Such methods are of general

applicability in that only a numerical representation of the refractive index is required- experimental values may be read off a graph, for example. Care however has to be observed with respect to two important considerations, namely numerical accuracy and calculational efficiency. For both of these problems, the tests by which the given algorithms are judged are index profiles for which the trajectories are known exactly and which represent demanding tasks- an example might be one which caused the rays to deviate considerably from paraxial conditions, such as obtained with lenses of low f-number.

Moore suggested a procedure which involved taking power-series expansions of both the refractive index and ray-position functions. He verified the procedure by comparing the results with Sands' third-order aberration theory [7]. His definition of accuracy was that the computed aberrations of analytically perfect (ie aberration-free) designs be no more than 10^{-6} times the Gaussian focal length, the aberrations being then of course artefacts of the algorithm. Moore also compared his results with the method developed by Montagnino [8]. Moore found excellent agreement with both Sands and Montagnino. In a later paper [9] he admitted that Montagnino's work was of more general validity in that it was not restricted to cases of spherical symmetry.

Montagnino's method formed the basis of a later work by Southwell [10] on inhomogeneous overlay lenses in planar guided wave optics of the type considered herein. The method will be treated here in some detail as it is the one implemented by the author in his work. The development is substantially as given by Southwell.

4.4.2 The Southwell-Montagnino Method

The tracing of a single ray from the entrance surface to the exit surface of an inhomogeneous lens possessed of circular symmetry is considered. As the lens is taken to be of the planar inhomogeneous type, the trajectory of the ray is a plane curve lying in the plane of the ambient waveguide.

The method employs a Taylor expansion of both the ray position vector $\underline{r}$ and its derivative, the ray direction vector $\underline{t}$. The refractive index function is specified either numerically (as a set of discrete data points) or as an analytical function of the radial coordinate r. The index derivatives may also be specified in analytical form but are generally evaluated using a numerical method. Both the index and its derivatives are assumed to be continuous.

The ray position and direction vectors take the forms

$$\underline{r} = x\mathbf{i} + y\mathbf{j}$$
$$\text{and } \underline{t} = \alpha\mathbf{i} + \beta\mathbf{j} \qquad (4.21)$$

in a two-dimensional coordinate system. $\underline{t}$ is a unit vector tangential to the ray and thus may be represented, as in (4.4), by

$$\underline{t} = \frac{d\underline{r}}{ds} \qquad (4.22)$$

where s is the scalar, physical distance along the ray curve. The components α and β are the direction cosines of the ray.

Once the initial vectors have been specified at the lens entrance an iterated extrapolation procedure is used to propagate the ray from point to point, with the four coordinates that describe the ray being calculated from the values taken after the previous iteration, and the refractive index function and its derivatives. The position vector is expanded in a Taylor series about the known point specified by the path length at the point, s_o :

$$\underline{r}(s_o + h) = \underline{r}(s_o) + h\frac{d\underline{r}}{ds} + \frac{h^2}{2!}\frac{d^2\underline{r}}{ds^2} + \frac{h^3}{3!}\frac{d^3\underline{r}}{ds^3} + \ldots \qquad (4.23)$$

where h is a small incremental distance along the ray path. As already observed, $d\underline{r}/ds$ is the direction vector $\underline{t}$. The second derivative is the curvature vector of the ray, as specified by equation (4.10).

$\underline{t}$ may be expanded in a fashion similar to $\underline{r}$;

$$\underline{t}(s_o + h) = \underline{t}(s_o) + h\underline{K} + \frac{h^2}{2!}\frac{d\underline{K}}{ds} + \ldots \qquad (4.24)$$

The problem becomes one of establishing the numerical values of the required derivatives $\underline{K}$ and $d\underline{K}/ds$ at the point $\underline{r}(s_o)$, with the direction vector $\underline{t}(s_o)$ being presumed known, either at the entrance surface of the lens or from a previous Taylor's expansion for a point inside the lens.

Explicit expressions for $\underline{K}$ and $d\underline{K}/ds$ in circularly symmetric media were given in equations (4.16) and (4.18) respectively. These equations require a knowledge of the refractive index function and its first and second derivative (with respect to the radial distance r from the centre of the lens) at the point $\underline{r}(s_o)$. In certain circumstances all three of these functions may be specified analytically. In chapter two, for example, a polynomial expression was obtained that approximated closely the refractive index function for the particular type of inhomogeneous lens known as the Luneburg lens, which has perfect focusing properties. The expression is repeated here:

$$n(r) = \exp(w(\rho,F)) \qquad (4.25)$$

$$\text{where } w(\rho,F) = P_1(1-nr)^{1/2} + P_2(1-nr)^{3/2} + P_3(1-nr)^{5/2} + P_4(1-nr)^{7/2} + P_5(1-nr)^{9/2} \qquad (4.26)$$

with $\rho = nr$ and F being the focal length of the normalized lens.

The first derivative of the index with respect to the radial coordinate r may be found for the Luneburg lens profile specified by (4.25) using implicit differentiation. Differentiating both sides of (4.25) and dropping the use of ρ and F:

$$\frac{dn}{dr} = \exp(w)\frac{dw}{dr}$$

$$= n\frac{dw}{dr} \qquad (4.27)$$

dw/dr may be found by differentiating both sides of (4.26):

$$\frac{dw}{dr} = \frac{-P_1}{2}(1-nr)^{-1/2}\left(n + r\frac{dn}{dr}\right)$$

$$\frac{-3P_2}{2}(1-nr)^{1/2}\left(n + r\frac{dn}{dr}\right)$$

$$- \ldots \text{ etc}$$

$$= -\frac{\left(n + r\dfrac{dn}{dr}\right)G}{2(1-nr)^{1/2}} \qquad (4.28)$$

where:

$$G = P_1 + 3P_2(1-nr) + 5P_3(1-nr)^2 + 7P_4(1-nr)^3 + 9P_5(1-nr)^4 \qquad (4.29)$$

Substitution of (4.28) into (4.27) yields an explicit analytical formula for dn/dr:

$$\frac{dn}{dr} = \frac{-n^2 G}{(2(1-nr)^{1/2} + nrG)} \qquad (4.30)$$

In order to obtain the second derivative, both sides of (4.27) are differentiated:

$$\frac{d^2n}{dr^2} = n\frac{d^2w}{dr^2} + \frac{1}{n}\left(\frac{dn}{dr}\right)^2 \qquad (4.31)$$

Next, both sides of (4.28) are differentiated:

$$\frac{d^2w}{dr^2} = \frac{-\left(n + r\frac{dn}{dr}\right)^2 H}{4(1-nr)^{3/2}} - \frac{G}{2(1-nr)^{1/2}}\left(r\frac{d^2n}{dr^2} + 2\frac{dn}{dr}\right) \qquad (4.32)$$

where:

$$H = -P_1 + 3P_2(1-nr) + 15P_3(1-nr)^2 + 35P_4(1-nr)^3 + 63P_5(1-nr)^4 \qquad (4.33)$$

Substituting (4.32) into (4.31) and collecting the components of d^2n/dr^2 on the left hand side gives:

$$\frac{d^2n}{dr^2}\left(1 + \frac{nrG}{2(1-nr)^{1/2}}\right) = -n\left(\frac{\left(n + r\frac{dn}{dr}\right)^2 H}{4(1-nr)^{3/2}} + \frac{G}{(1-nr)^{1/2}}\frac{dn}{dr}\right) + \frac{1}{n}\left(\frac{dn}{dr}\right)^2 \qquad (4.34)$$

which yields:

$$\frac{d^2n}{dr^2} = \frac{-n\left(\frac{\left(n + r\frac{dn}{dr}\right)^2 H}{4(1-nr)} + G\frac{dn}{dr}\right) + \frac{(1-nr)^{1/2}}{n}\left(\frac{dn}{dr}\right)^2}{(1-nr)^{1/2} + nrG/2} \qquad (4.35)$$

This expression is considerably different from the one quoted by Southwell in his paper on Luneburg lenses [11].[1] In fact if Southwell's expression is used in ray-tracing, the algorithm fails. The validity of the above expression may be checked by determining whether or not it yields the correct (known) value at the edge of the normalized lens, where n and r = 1. Substituting

[1] see also Addenda, Note (1) p 215b

these values of n and r into (4.35) gives:

$$\left.\frac{d^2n}{dr^2}\right|_{edge} = -2 \qquad (4.36)$$

which is indeed the correct value [12]. Southwell's expression, however, gives the value of the second derivative at the lens edge as $(2-2/P_1)$ which implies a very small radius of curvature for the ray.

The availability of analytical expressions for the refractive index and its derivatives is very important for the accuracy of ray-tracing calculations. In some cases, (where fabricated profiles have been measured, for example), no explicit formulae are available and the required derivatives must be calculated using finite difference methods [13], viz

$$\frac{dn}{dr} = \frac{n(t+\Delta t) - n(t-\Delta t)}{2\Delta t} \qquad (4.37)$$

and

$$\frac{d^2n}{dr^2} = \frac{n(t+\Delta t) - 2n(t) + n(t-\Delta t)}{(\Delta t)^2} \qquad (4.38)$$

The thickness function t(r) of the inhomogeneous overlay lens is assumed to be known from either measurements on the fabricated profile or from theoretical predictions such as those discussed in chapter three.

The Taylor expansions given in equations (4.23) and (4.24) can now be used to predict the subsequent ray position $\underline{r}$ and direction $\underline{t}$. The procedure is continued until the exit surface has been crossed. The test for determining whether or not the exit surface has been crossed is simply

If $r < R$ continue extrapolation procedure;
if $r \geqslant R$ exit surface of lens has been crossed.

R is the maximum radius of the lens.

4.4.3 Exact point of crossing

After the ray has crossed the exit surface, the actual point of crossing and the direction cosines at that point are only imprecisely determined, and certainly not to the accuracy required of the algorithm which must yield errors of less than the diffraction limit. The accuracy could be improved by simply reducing the path increment h to a value as small as desired but this would greatly reduce the computational efficiency of the algorithm. Instead, an interpolation method is used to determine accurately the exit boundary location of the ray. A function F is defined, such that:

$$F = R^2 - \underline{r}.\underline{r}$$

$$= R^2 - r^2 \qquad (4.39)$$

with the origin of coordinates being taken at the centre of the lens. The function F represents the error with which the above algorithm is able to find the correct exit point for a given ray.

The penultimate point calculated inside the lens is taken as the base point. The ultimate point of intercept may then be considered as a function of a variable ray path increment, h. As such, F becomes functionally dependent on h according to the Taylor expansion given in equation (4.23). Newton's method is used to find the value h that makes F(h) as close as desired to zero, viz:

$$h_{i+1} = h_i - \frac{F}{(\partial F/\partial h)_{h=hi}} \qquad (4.40)$$

To use the method the partial derivative $\partial F/\partial h$ must be obtained. It is sufficient to return to equation (4.39) and to observe that

$$\frac{\partial F}{\partial h} = -2r.\frac{d\underline{r}}{dh} = -2\underline{r}.\underline{t} \qquad (4.41)$$

Since the partial derivative of F with respect to h is desired, the other variables in the Taylor expansion of (4.23) ($\underline{r}$ and $\underline{t}$) must be held constant. In consequence, the partial derivative itself remains constant throughout the Newton iteration since it is functionally specified through (4.41).

4.4.4 Interpolation procedure to allow larger stepsizes

The incremental arclength step chosen determines both the computational efficiency and the final accuracy of the method. Too few steps leads to inaccurate results, whilst too many may mean that the trace may take up too much computer time. Richardson's extrapolation method [13] is used to achieve a high degree of accuracy in reasonable time.

Each ray is traced a total of three times with each successive trace reducing the stepsize by a factor of two. It may then be shown that the formula:

$$y = \frac{y_1}{3} - 2y_2 + \frac{8y_3}{3} \qquad (4.42)$$

accurately determines the ray exit height.

The algorithm described above can be implemented with only a moderate amount of computer code by making judicious use of subroutines. It copes with lenses of very small f-numbers (down to f/0.5) with excellent results.

4.5 Tracing rays through geodesic lenses

4.5.1 Southwell's method

Tracing rays through geodesic lenses requires a different technique from that used for overlay lenses. Southwell [14] has produced a suitable algorithm.

A curve z(r) is assumed to generate the profile when rotated about a central axis, as shown in Figure [4.2]. The points on the curve are assumed to be specified by either an analytic function, or by some other means. Fermat's principle of minimum optical path is the basis for calculating ray trajectories.

The optical path length integral OP is given by:

$$OP = \int_A^B ndL \qquad (4.43)$$

where A and B are separate points on a ray- trajectory on the surface of the lens, n is the refractive index and L is the physical length along the trajectory. n is assumed to be constant over the surface of a geodesic lens, so that the variational displacements involve only the physical path. Fermat's principle of shortest path is then given by:

$$\delta(OP) = \delta\int_A^B dL = 0 \qquad (4.44)$$

where δ indicates the variation. The principle states that the path-length along the curve followed by the ray is a minimum with respect to other curves in the vicinity. The paths so described are termed geodesics. A constraint is that the ray remain on the surface of the lens. The geometry of Figure [4.3] shows that:

$$dL = (dS^2 + r^2 d\phi^2)^{1/2} \qquad (4.45)$$

where r,ϕ are polar coordinates and S is orthogonal to $rd\phi$. The variable S may be put in terms of the generating curve z(r), since the elements dS and $rd\phi$ are orthogonal:

$$dS^2 = dr^2 + dz^2 \qquad (4.46)$$

Thus equation (4.44) becomes:

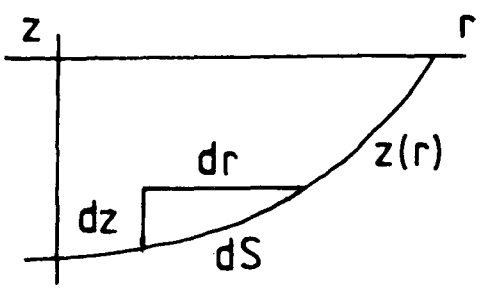


FIGURE [4.2] Geodesic lens generating curve z(r)

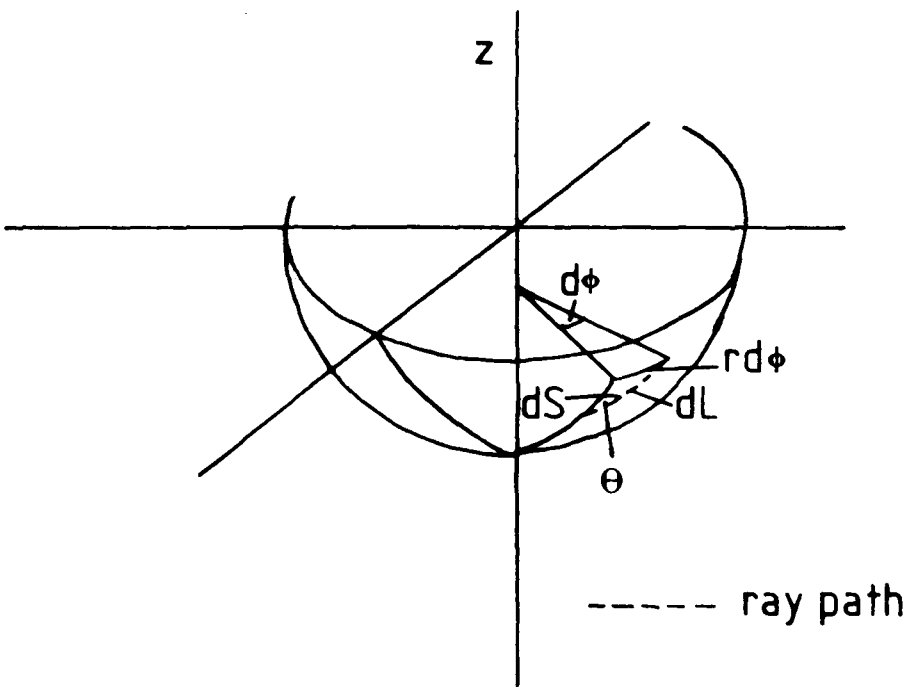


FIGURE [4.3] Geodesic lens depression

$$\int_A^B \left[1 + \left(\frac{dz}{dr}\right)^2 + r^2\left(\frac{d\phi}{dr}\right)^2\right]^{1/2} dr = 0 \qquad (4.47)$$

The only dependent variable is ϕ since z(r) is assumed known. The solution is given by the solution of the Euler equation [1] :

$$\frac{d}{dr}\left(\frac{\partial F}{\partial \phi'}\right) + \frac{\partial F}{\partial \phi} = 0 \qquad (4.48)$$

where F is the integrand of equation (4.47) and:

$$\phi' = \frac{d\phi}{dr} \qquad (4.49)$$

From equation (4.47) it is obvious by inspection that F does not depend explicitly on ϕ but rather on the first derivative ϕ'. Thus the second term on the left-hand side of equation (4.48) vanishes. Integrating the remaining terms on the left-hand and right-hand sides yields:

$$\frac{\partial F}{\partial \phi'} - c = 0 \qquad (4.50)$$

where c is a constant of integration over the path. Thus, from the definition of F,

$$\frac{r^2}{\left[1 + \left(\frac{dz}{dr}\right)^2 + r^2\left(\frac{d\phi}{dr}\right)^2\right]^{1/2}} \frac{d\phi}{dr} = c \qquad (4.51)$$

By using equations (4.45) and (4.46), equation (4.51) becomes:

$$r^2 \frac{d\phi}{dL} = c \qquad (4.52)$$

If the angle between the meridional curve on the surface (defined by the angle ϕ) and the ray-path is called Θ, then, on inspection of the infinitesimals of Figure [4.3]:

$$\sin\Theta = \frac{r d\phi}{dL} , \qquad (4.53)$$

and hence equation (4.52) becomes:

$$r\sin\Theta = c \qquad (4.54)$$

The ray enters the lens at the edge where the radial distance from the centre is r_1 makes the angle Θ_1 with respect to the meridional line at that point. Hence the constant c is determined by:

$$c = r_1 \sin\Theta_1 \qquad (4.55)$$

The exit-angle that the ray makes with the meridional line as the ray leaves the lens must also be Θ_1, since r again equals r_1 at that point. The angle Θ increases monotonically from the value Θ_1 at the entrance and reaches a maximum value of $\pi/2$ at the point where r reaches a minimum:

$$r_{min} = c \qquad (4.56)$$

Thereafter, Θ decreases monotonically to Θ_1 at the exit. Eq (4.56) shows that the constant h is the distance of closest approach to the axis of rotational symmetry.

The rays entering and leaving the lens may be characterised according to the geometry of Figure [4.4] which shows the projection of the ray-path onto the plane of the ambient waveguide. Two parameters define the ray uniquely, ie the position of the ray at the entrance is specified by the azimuthal angle ϕ_1 and the direction of the ray at that point is defined by Θ_1. The corresponding exit parameters are then ϕ_{out} and Θ_1. The values Θ_1 and ϕ_1 are generally set at the commencement of the

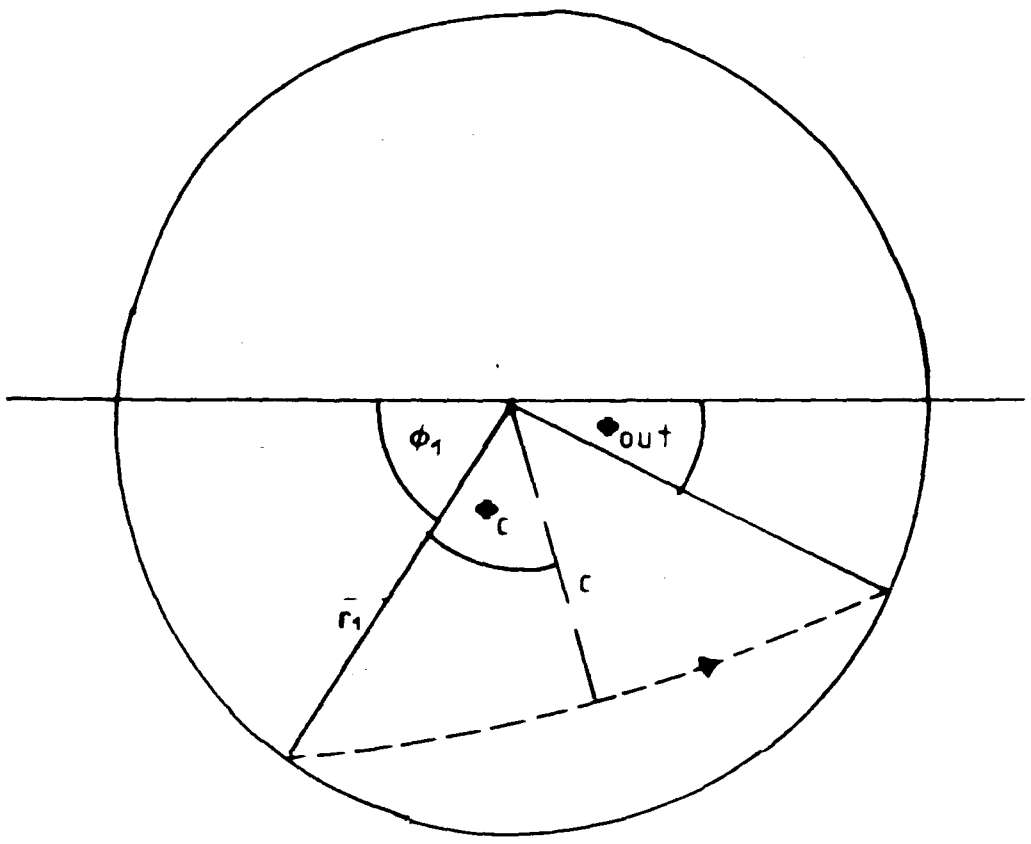


FIGURE [4.4] Projection of the ray path onto the waveguide plane

ray-trace. Only ϕ_{out} then remains to be determined.

Squaring both sides of equation (4.51) allows a separation of variables to be obtained, and this then produces an integrable equation, viz:

$$d\phi = \frac{c(1 + z'(r)^2)^{1/2}}{r(r^2 - h^2)^{1/2}}\, dr \qquad (4.57)$$

Observing, from the geometry of Figure [4.4], that

$$\phi_{out} = \pi - \phi_1 - 2\phi_c \ , \qquad (4.58)$$

it is seen that ϕ varies from ϕ_1 to $\phi_1+\phi_c$ while r varies from r_1 to c. Thus limits may be placed on the integral obtained from equation (4.57):

$$\phi_c = \int_{\phi_1}^{\phi_1+\phi_c} d\phi = c \int_{r_1}^{c} \frac{(1 + z'(r)^2)}{r(r^2-h^2)^{1/2}}\, dr \qquad (4.59)$$

The upper limit of the integral produces a singularity in the integrand. The singularity is removed by integrating eq (4.59) by parts:

$$\phi_c = c \left[\frac{-2(r_1-c)^{1/2}(1+z'^2)^{1/2}}{r_1(r_1+c)^{1/2}} \right.$$

$$\left. -2 \int_{r_1}^{c} \frac{(r-c)^{1/2}(1+z'^2)^{1/2}}{r(r+c)^{1/2}} \left(\frac{z'z''}{(1+z'^2)} - \frac{1}{r} - \frac{1}{2(r+c)} \right) dr \right]$$

(4.60)

In general, the integral of equation (4.60) must be evaluated numerically. The generating curve z(r) is specified at a large number of points (typically 1000) and the derivatives are

calculated using finite difference methods. The integral is then calculated using Gaussian quadrature methods. The tracing of the ray through the lens is completed with the calculation of the integral.

4.6 Results

4.6.1 Introduction

The ray-tracing algorithms developed in the previous section will be used to show how the behaviour of real lenses is likely to deviate from that of perfect lenses. In the case of overlay Luneburg lenses, the achievable, approximate lens-profile calculated in chapter three will be compared with the perfect Luneburg overlay lens. In the case of geodesic lenses, the behaviour of the perfect profile will be compared with:

(a) that of a spherical depression lens for which an analytical geometrical-optics theory is available, and

(b) that of a profile designed to approximate closely the perfect geodesic lens when fabricated using computer-controlled single-point diamond-turning techniques [15].

The behaviour of a long-focal-length spherical geodesic lens with a toroidal rounded-edge will also be investigated. On-axis propagation will be modelled throughout, so that only spherical aberrations need be considered.

4.6.2 Ray-tracing through an overlay Luneburg lens and a possible approximation

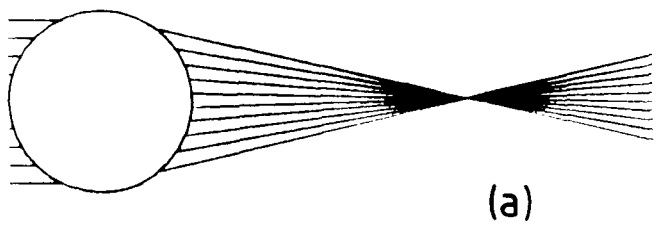
Figure [4.5] (a) shows a computed ray-trace through a perfect overlay Luneburg lens having a full-aperture f-number of f/2. The lens is shown operating at a reduced aperture of f/2.2 . All the incident rays are seen to converge at one axial point. This ray-trace was carried out using the quasi-analytical expressions for the refractive index and its derivatives

developed earlier in this chapter. Figure [4.5] (b) shows a computed ray-trace through the approximate overlay lens-profile calculated in chapter three. The approximate profile is one that is amenable to fabrication. The approximate profile is shown operating at a reduced f-number of f/4.2 . Attempts to trace rays through the approximate lens at wider apertures result in failure, due to non-convergent solutions for the ray-trajectories within the lens. Such non-convergence indicates that the refractive index gradients within the lens are large at the lens margins. Figure [4.5] (c) shows a blow-up of the focal region of the approximate lens. The approximate lens is under-corrected with respect to the perfect lens, since the marginal rays are seen to focus at points closer to the lens than the paraxial rays. Since no analytical expressions were available for the refractive index profile of the approximate lens, this ray-trace was carried out by specifying the index profile at 29 points and using linear interpolation to calculate the index between these points. Numerical finite-difference methods were used to calculate the derivatives of the refractive index and therefore numerical artefacts can creep into the calculations, as demonstrated by some slight asymmetries in Figure [4.5] (c).

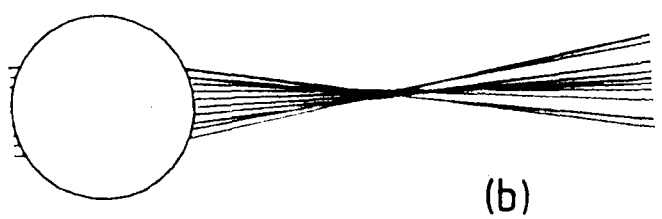
In order to interpret the data obtained from these traces, a simple data-reduction method is employed. At each point z of interest along the optical axis of propagation, the transverse distance y_i of each ray i is measured. These distances are squared and summed over all the rays and a root-mean-square value is calculated, viz:

$$f(z)_{rms} = \sum_{i=1}^{N} \frac{y_i^2}{N} \qquad (4.61)$$

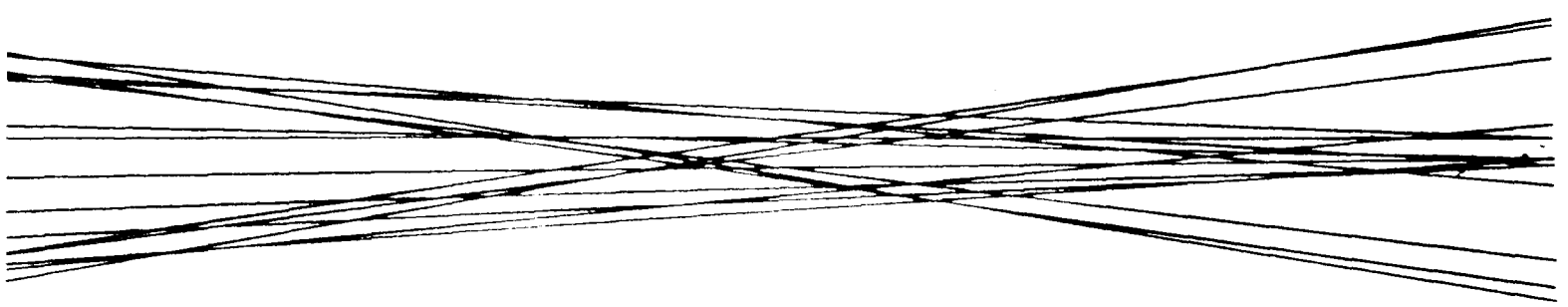
A similar formula over two transverse dimensions is used in bulk optics. The quantity is often called the 'radius of gyration' , in analogy with a quantity similarly defined in mechanics.



(a)



(b)



(c)

FIGURE [4,5] RAY-TRACES THROUGH OVERLAY LENSES

The function f(z) is plotted against the axial distance z in Figure [4.6] for the perfect lens and its approximation. The minimum of the function f(z) may be taken as the 'best' focus. In bulk optics, the best focus is generally known as the 'circle of least confusion'. The minimum of f(z) associated with the approximate profile is located at an axial distance of 3.2mm (the radius of the lens has been scaled, for convenience, to 1.0mm, the scaling being unimportant in geometrical optics). In comparison, the perfect lens has a minimum at 4.0mm, the correct value. Once again the approximate profile is seen to be under-corrected with respect to the perfect lens. Furthermore, the minimum associated with the approximate profile is not zero, indicating that the image is geometrically spread and is not, therefore, diffraction-limited.

The point at which a given ray intercepts the optical axis is the focus for that ray, and the difference between it and the paraxial or design focus is the longitudinal component of spherical aberration [16]. The longitudinal component of spherical aberration is plotted as a function of the position of the ray in the entrance manifold in Figure [4.7] for both the perfect and approximate lens. The perfect lens displays no spherical aberration over the entire aperture. The approximate profile, however, displays considerable spherical aberration. The paraxial rays deviate from the focus of the perfect lens, in a positive direction (away from the lens), by 40 μm. Rays situated at 0.4 mm ($f/_{eff}$ = f/5) in the entrance manifold display the greatest amount of longitudinal spherical aberration, of approximately 900 μm in the negative direction (towards the lens). Thus paraxial rays passing through the approximate lens-profile are slightly over-corrected with respect to the perfect lens, whilst the rays passing through most of the aperture are considerably under-corrected. The behaviour of the longitudinal spherical aberration function for the approximate lens is quite complicated, suggesting that simple third-order spherical aberration is not the only contribution to the errors in the image, and that substantial higher orders are present. A good treatment of different orders of geometrical aberration and their

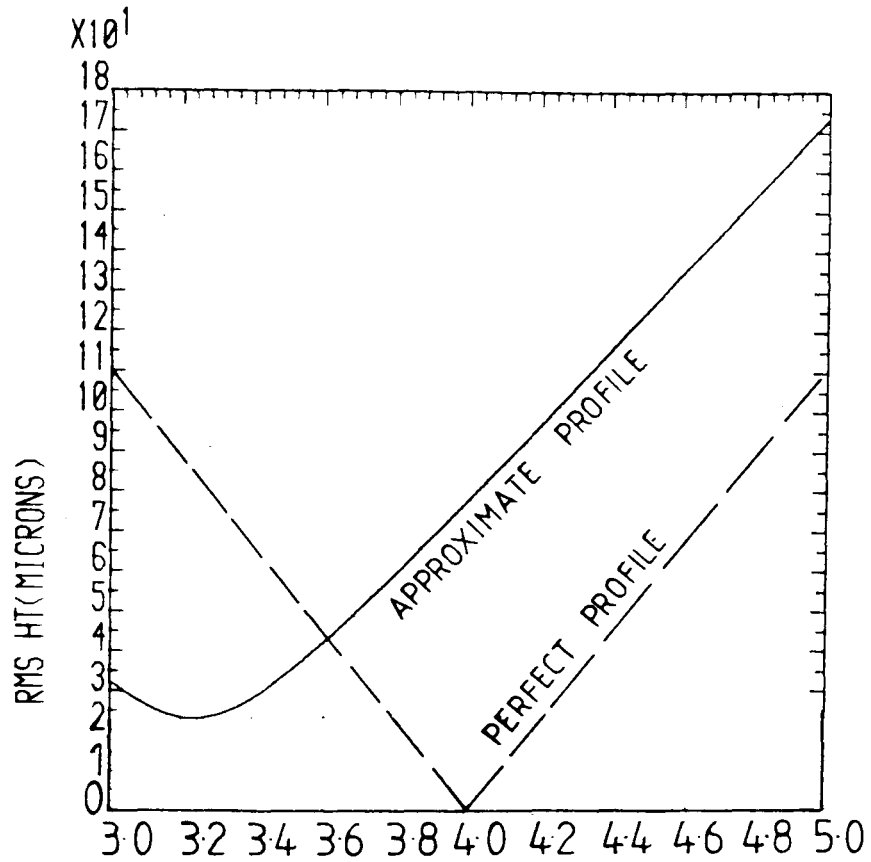


FIGURE [4,6] VARIATION OF GEOMETRICAL SPOT-RADIUS WITH AXIAL DISTANCE FOR TWO OVERLAY LENS PROFILES

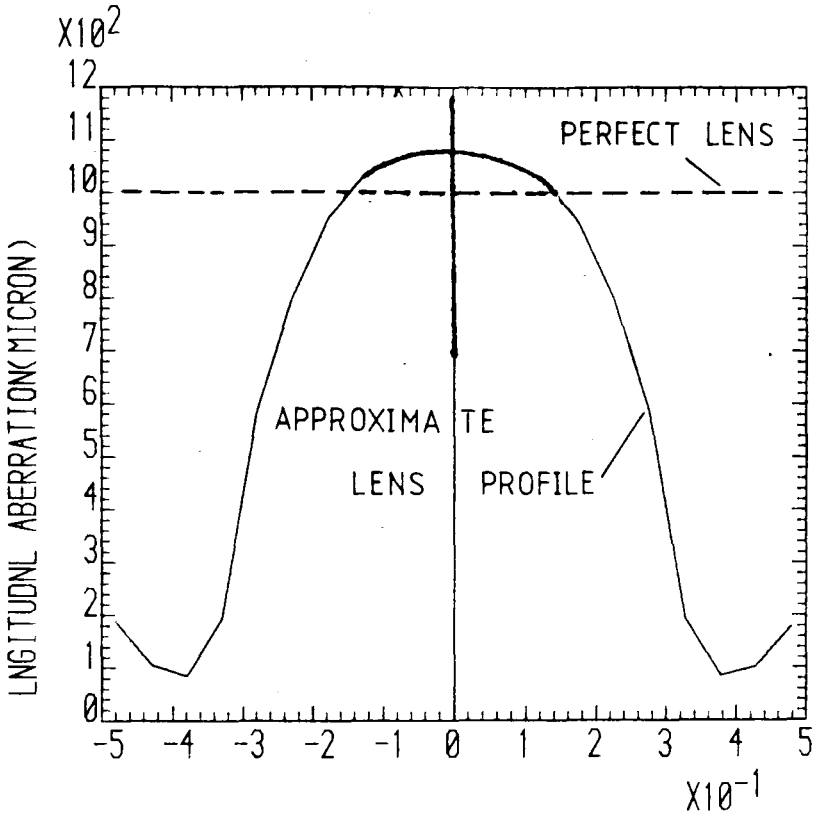


FIGURE [4,7] PLOTS OF LONGITUDINAL COMPONENT OF SPHERICAL ABERRATION (L-SA)

effect on the focal image is given by Nijboer [17].

At a given axial image distance z, the transverse distance of the ray from the axis may be recorded as a function of the position of the ray in the entrance manifold. Such functions represent the lateral component of spherical aberration. The lateral spherical aberration functions for the perfect and approximate profiles are shown in Figure [4.8] . The functions shown are calculated at the points of best focus for each lens. Once again, the perfect lens displays no aberration, whereas the approximate profile displays quite severe aberrations. Rays situated at 0.24 mm from the axis in the entrance manifold ($f/_{eff}$ = f/8.3) are situated at 40 μm from the axis in the best image. The area under the curve of Figure [4.8] can be calculated using Simpson's rule. This area, divided by the width of the entrance manifold, is a good measure of the geometrical spread, or 'blur-spot'. The blur-spot of the approximate profile is 48 μm in diameter. In physical optics, the diffraction-limit of the lens is given approximately by:

$$\text{diffraction-limit} = \text{f-number}\ \frac{\lambda}{n} \qquad (4.62)$$

where λ is the operating wavelength. For an f/4.2 lens and a wavelength of 0.633 μm, in a medium with ambient refractive index of 2.2 (which is representative of possible operating conditions), the diffraction-limit turns out to be 1.2 μm. Clearly, the approximate profile is not diffraction-limited.

4.6.3 Discussion

The performance of the approximate lens is easily interpreted by looking at the refractive index profile, in comparison with that of the perfect lens (chapter three, Figure [3.9]). Over most of the available aperture the refractive index of the approximate lens is greater than that of the perfect lens. Thus a more powerful focusing action action would be expected from the approximate lens. A focal-shift towards the lens would be expected as a result. Furthermore, the approximate profile is

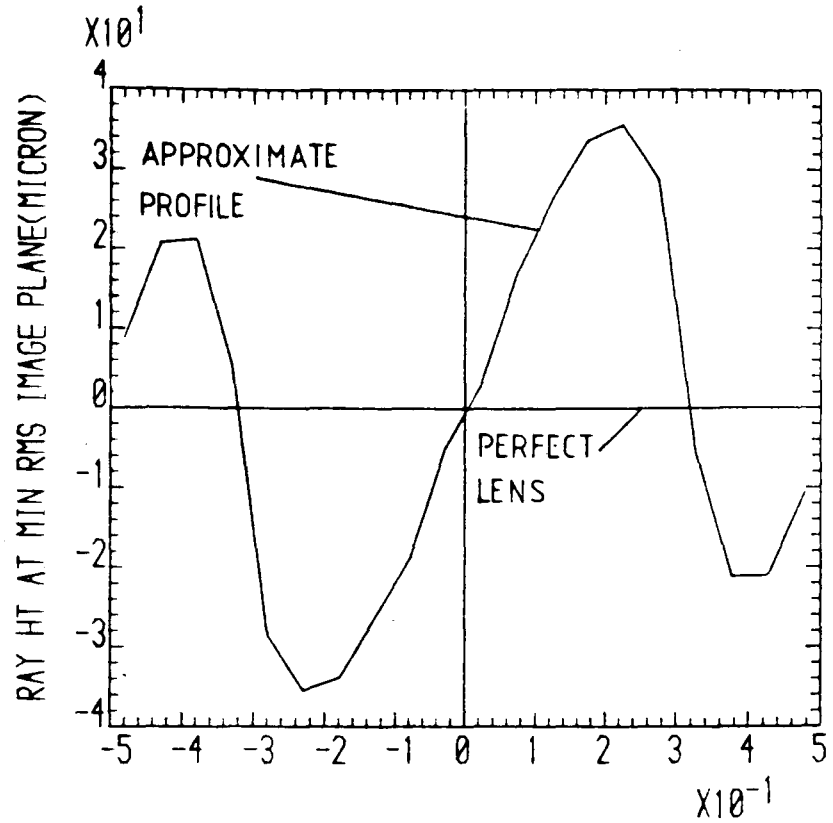


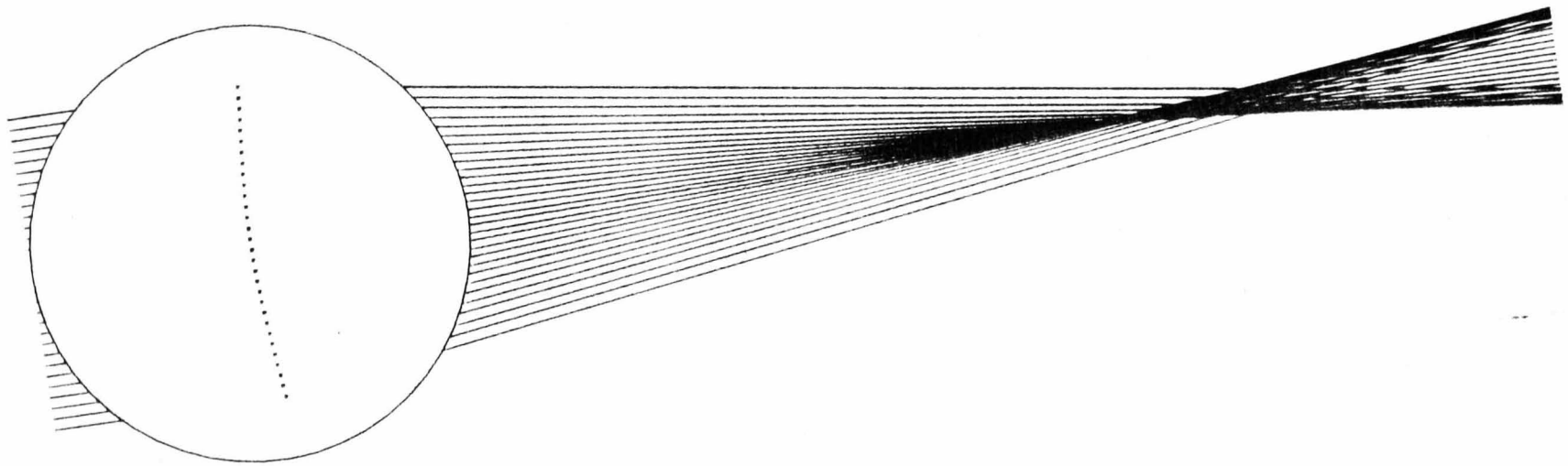
FIGURE [4.8] PLOTS OF TRANSVERSE COMPONENT OF SPHERICAL ABERRATION (T-SA)

relatively flat-topped, in comparison with the perfect profile, over a considerable part of the aperture. A flat profile implies that the approximate lens would behave somewhat similarly to a bulk-optics cylindrical lens with surfaces based on circular geometries. Thus spherical aberration would be expected.

These findings imply that the vacuum-evaporation shadow-masking method outlined in chapter three is not sufficiently sophisticated to produce good Luneburg overlay lenses. Other improvements to the method will be suggested in chapter seven of this thesis, but one particular improvement that could be carried out would be to operate the ray-tracing and shadow-masking computer programs conjointly. The criterion by which the approximate lens profile was chosen in chapter three was based on a least-squares fit to the physical profile of the ideal lens. An alternative criterion would be to assess the merits of the approximate profile purely on optical performance as measured by ray-tracing.

4.6.4 Ray-tracing through geodesic lenses

In Figure [4.9] ray-traces through a perfect geodesic lens and a spherical-depression geodesic lens are compared. The spherical-depression lens would be an unlikely candidate for fabrication, in reality, since the abrupt transition from the lens to the ambient waveguide would cause severe losses. The lens is, however, an interesting case from a conceptual point of view since its behaviour can be described analytically. Both the lenses of Figure [4.9] have an overall diameter of 10.0mm and a paraxial focal-length of 18.5mm, and are shown operating with an incident ray-manifold of width 7.4 mm , ie at an f-number of f/2.5 . The differences in optical behaviour are quite dramatic, and are further illustrated by the curves of longitudinal and transverse spherical aberration shown in Figures [4.10] and [4.11] . A theory of the aberrations produced by a spherical geodesic lens-depression has been given by Vahey [18]. For a lens with a radius of curvature C, a maximum depth H relative to the ambient waveguide and a paraxial focal-length f_o, the third-order



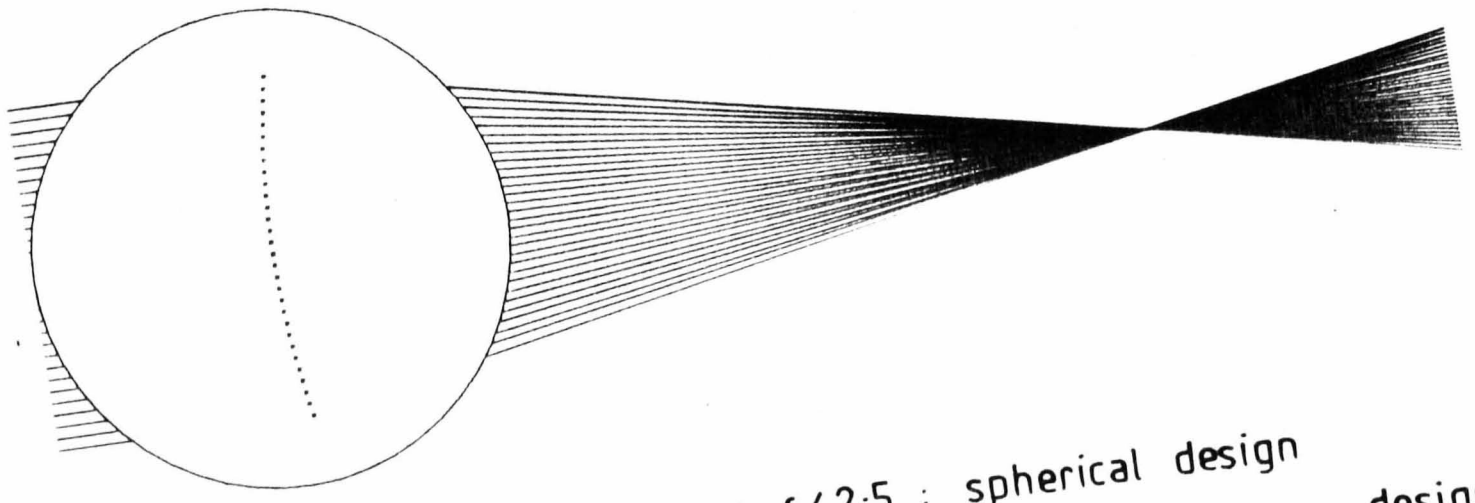


FIGURE [4.9] Geodesic lenses (a) f/2·5 : spherical design
(b) f/2·5 : perfect design

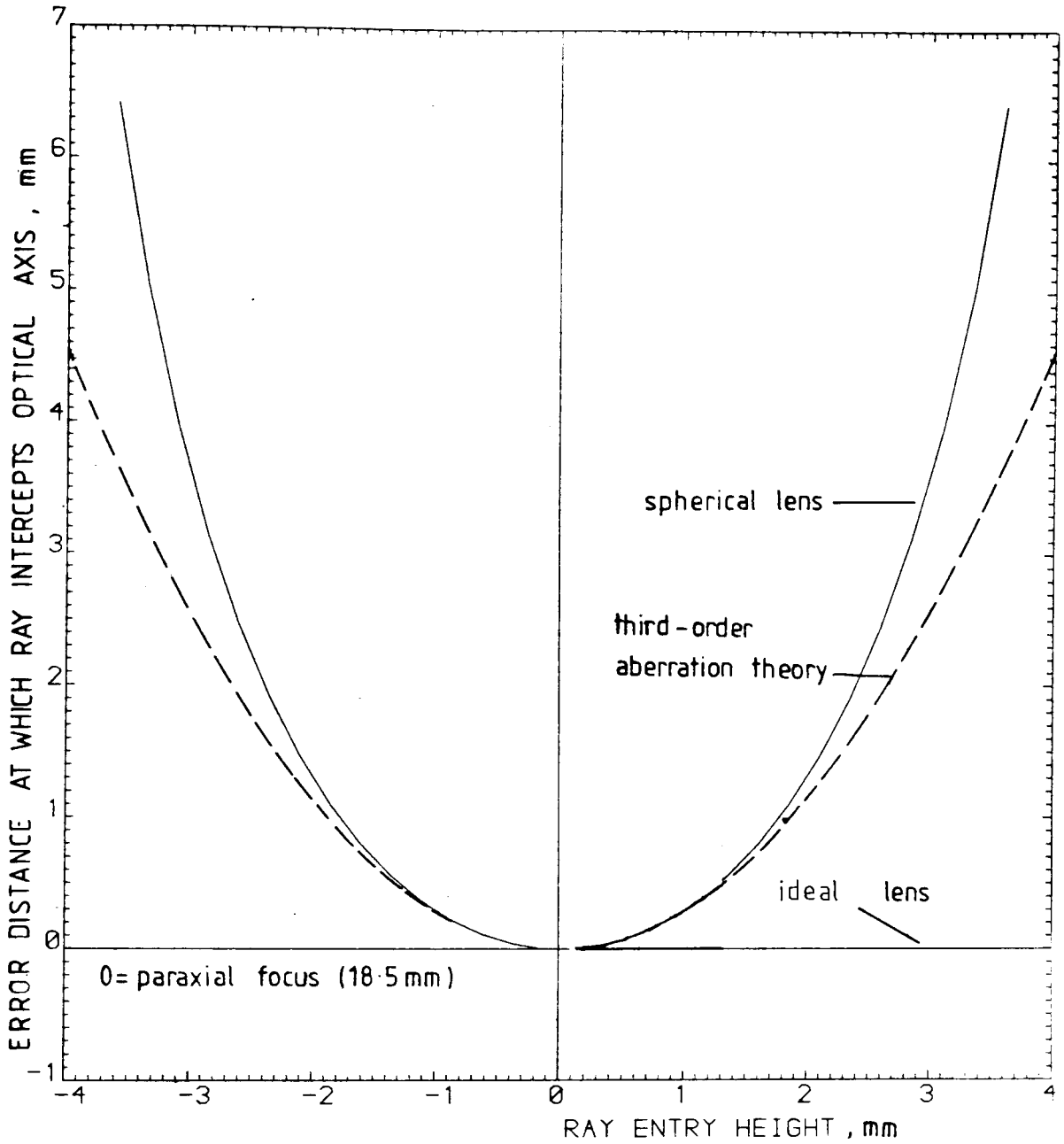


FIGURE [4,10] Curves of longitudinal component of spherical aberration. The ideal geodesic lens is aberration - free.

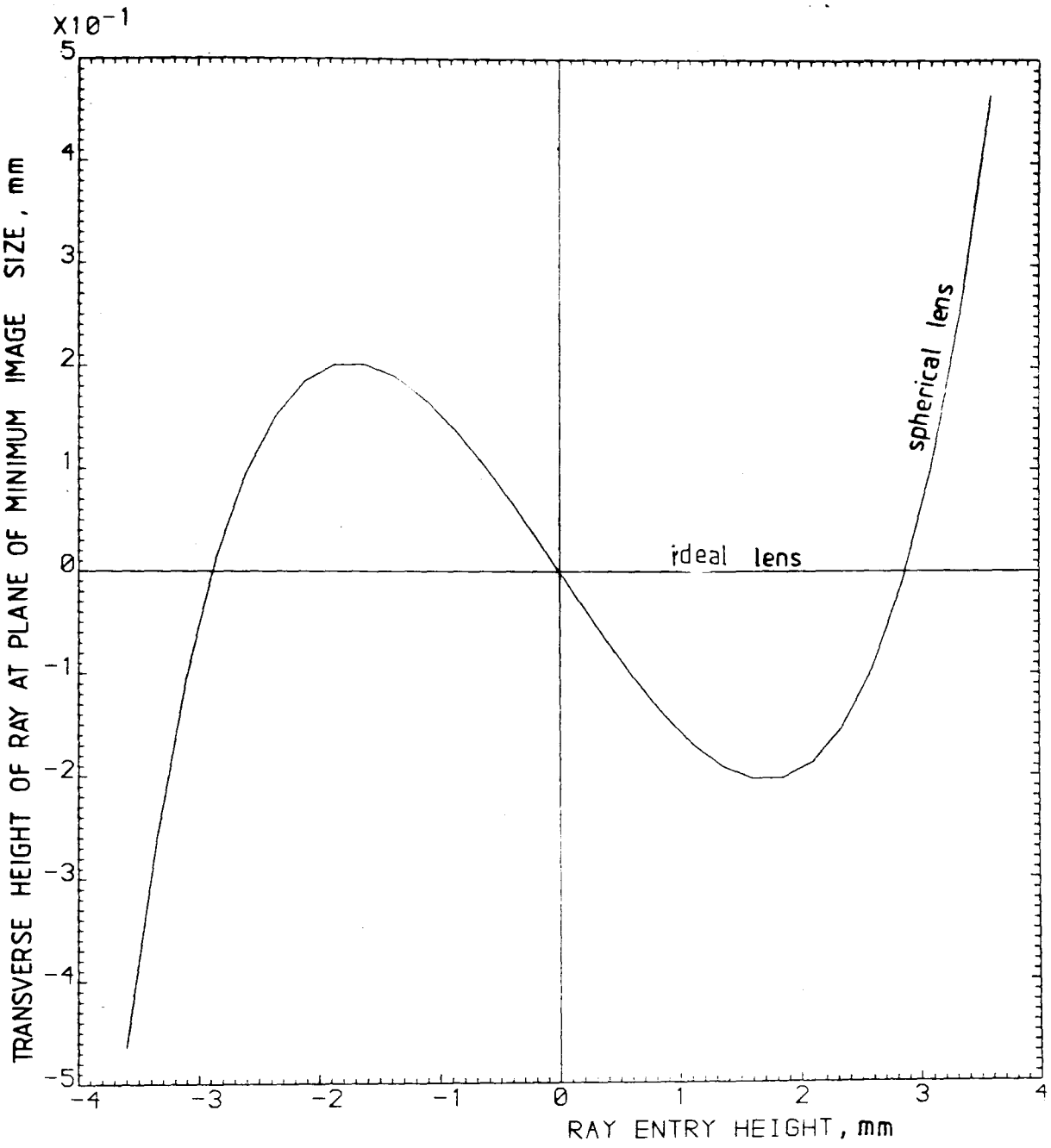


FIGURE [4,11] Curves of transverse component of spherical aberration.

spherical aberration coefficient is given by:

$$f_1 = \frac{(C/H - 1)^2 + 1}{8f_o} \qquad (4.63)$$

The focal length f(x) of a ray incident on the lens at height x is then given by, to third-order,

$$f(x) = f_o + f_1 x^2 \qquad (4.64)$$

The longitudinal component of third-order spherical aberration, L-SA, is then given by :

$$L\text{-}SA = f_1 x^2 \qquad (4.65)$$

For the spherical lens being considered, C=9.96 mm, H=1.346mm and f_o=18.5mm . A plot of the third-order component of longitudinal spherical aberration is shown together with the computed curve in Figure [4.10] . It is clear that the spherical-depression geodesic lens produces significant amounts of higher-order spherical aberration.

In order to fabricate practical geodesic lenses using CNC lathes, the lens profile is approximated by a series of circular arcs [15]. The procedure is shown in Figure [4.12] . The arc-fitting procedure fits a practical profile to the desired lens-profile to within 0.1 μm over most of the profile, except at the boundaries between the edge-rounding region, the ambient waveguide and the inner portion of the lens. At these points the deviations from the desired profile can be up to 2.5 μm . Figure [4.13] shows the transverse spherical aberration arising from one such arc-fitting procedure, where 22 arcs are used. The optical path through the lens is clearly perturbed in a complicated but deterministic fashion. The transverse spherical aberration resulting from the approximate profile is seen to be less than 1.4 μm over the aperture, and the geometrical blur-spot diameter is only 0.4 μm . The arc-fitting procedure would therefore appear

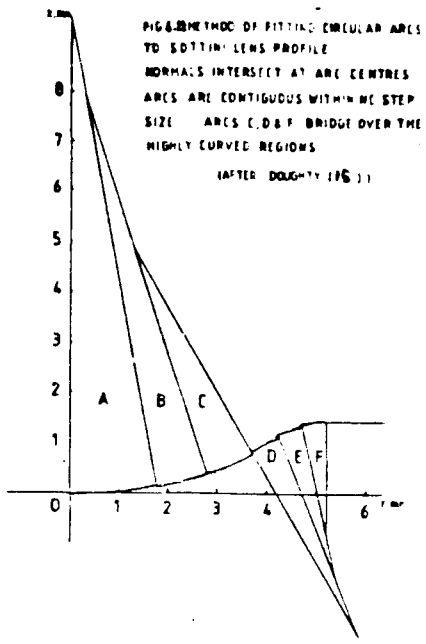


FIGURE [4,12] METHOD OF FITTING CIRCULAR ARCS TO IDEAL LENS

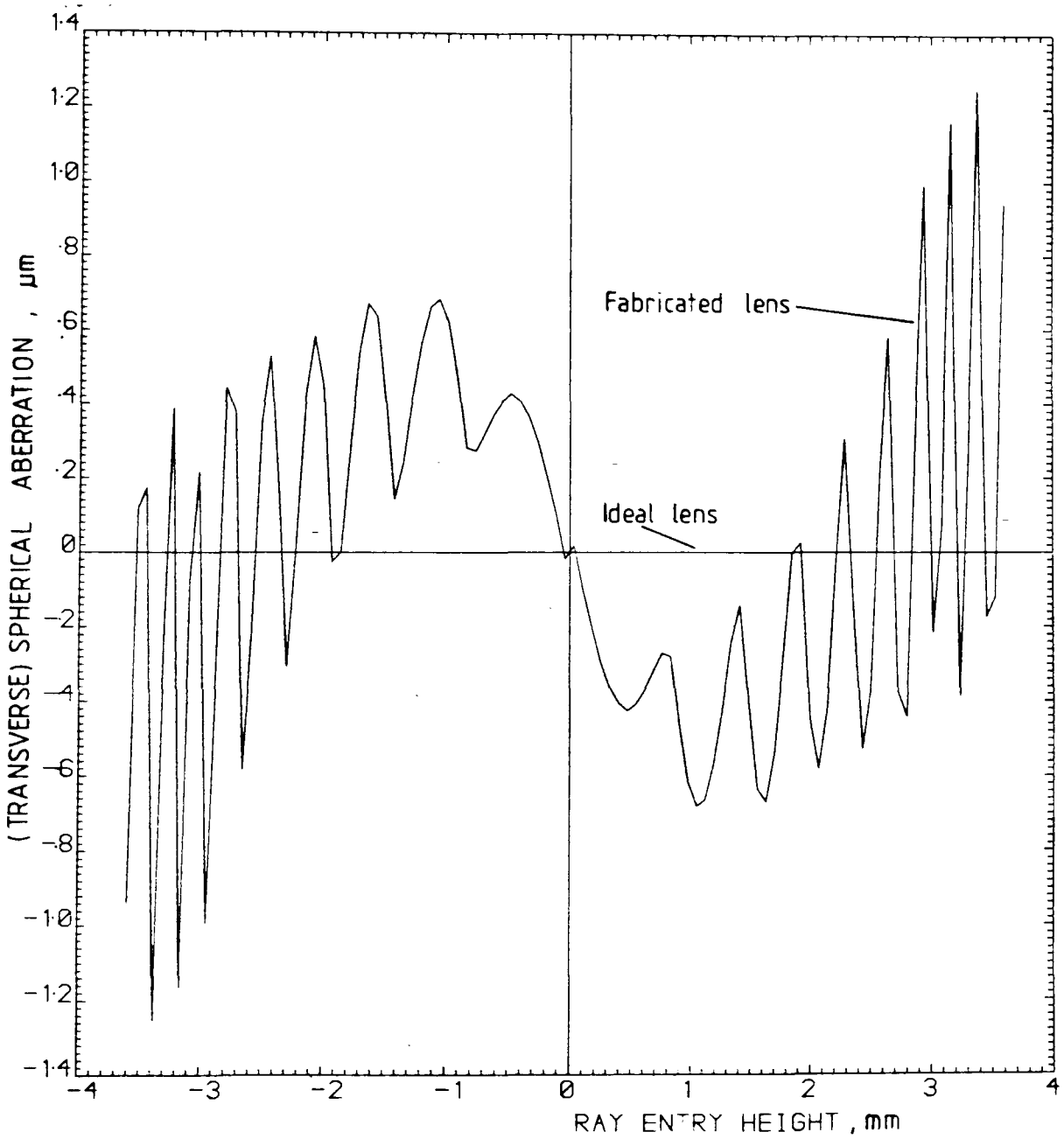


FIGURE [4,13] CURVES OF TRANSVERSE SPHERICAL ABERRATION

not to have greatly affected the lens properties. However, an inspection of the longitudinal spherical aberration reveals that the optical properties are adversely affected, as shown in Figure [4.14] . Again, complicated aberration behaviour is manifest, directly attributable to the perturbations of the ideal profile. The important point to note is that the rays intercept the optical axis within a band 25 μm to 55 μm short of the design focus, with the best image occuring at 34 μm from the design focus. Such a focal-shift could be unacceptable for high-resolution requirements. In order to minimise the shift, an alternative series of 20 arcs was fitted to the desired profile. The resultant longitudinal spherical aberration is shown in Figure [4.15] . The focal shift is now reduced to 3.0 μm, at the expense of some large aberrations at the edge of the useful aperture. Since the lenses are only expected to utilise approximately 4.4 mm (or 60%) of the useful aperture, these large aberrations are unimportant. The geometrical blur-spot radii arising from the 20-arc approximate profile is compared with that of the ideal in Figure [4.16] . The profile is seen to be capable of yielding diffraction-limited performance over input-beamwidths as large as 6.0mm, corresponding to f/3.1 operation. The performance is noticeably degraded for larger incident beamwidths.

Figure [4.17] shows a ray-trace through a geodesic lens with a spherical inner region and a toroidal edge-rounding region. The lens has a design focal length of 50.9 mm and an overall diameter of 10.462 mm. An input ray manifold of diameter 8.0 mm is shown incident on the lens, corresponding to an f-number of f/6.36 . The marginal rays clearly focus at points much further away from the lens than the paraxial focus. The image of least confusion is situated at 71.4 mm. The transverse spherical aberration at both the paraxial focal length and at the distance of least confusion is shown in Figure [4.18] . This graph shows how a shift in the point of observation of the light minimises the area under the transverse aberration curve, and the sharpest geometrical image thus obtained. In actual operation the input beamwidth would be only 3.0mm, corresponding to a speed of f/17. In this case the

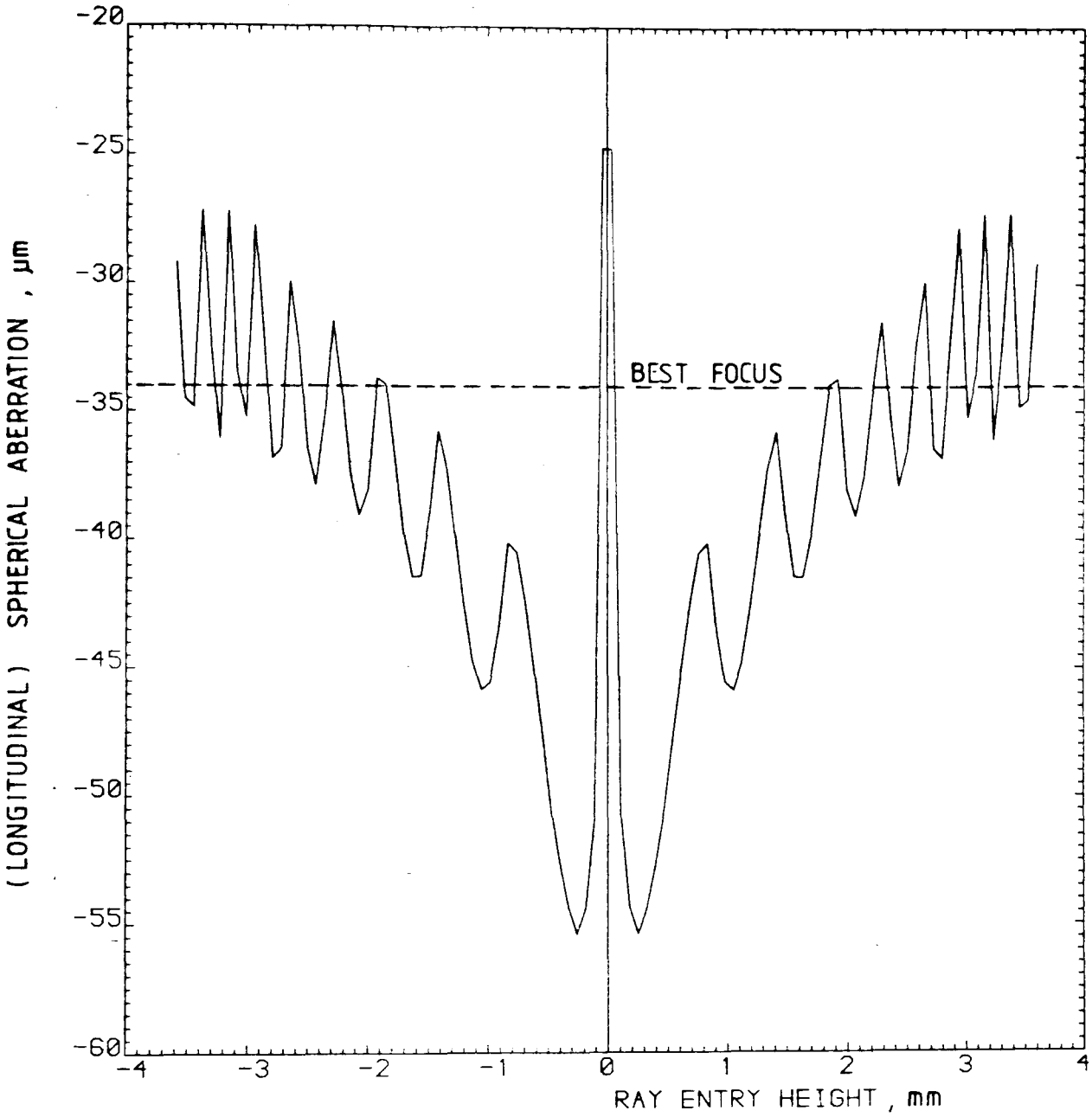
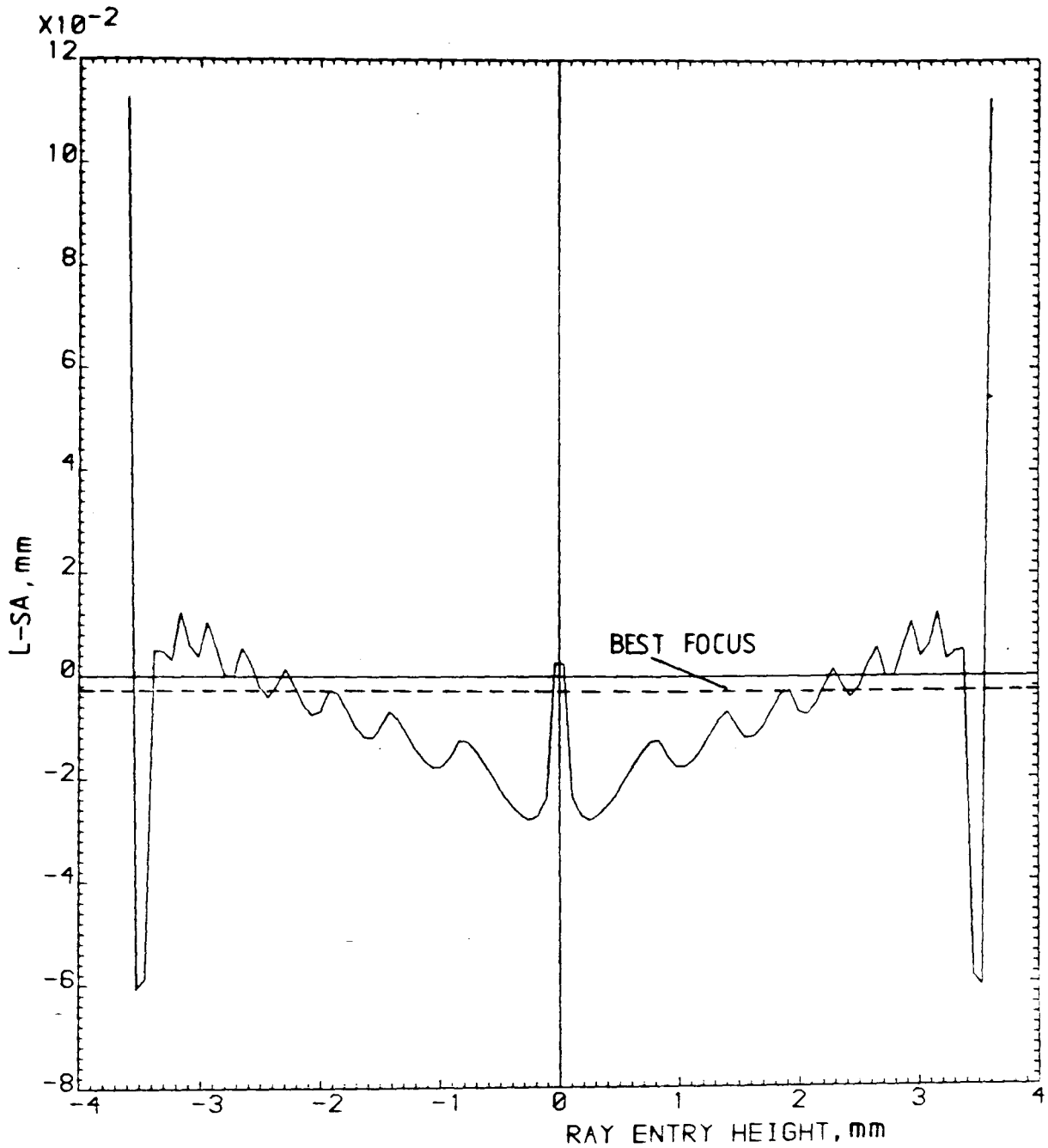
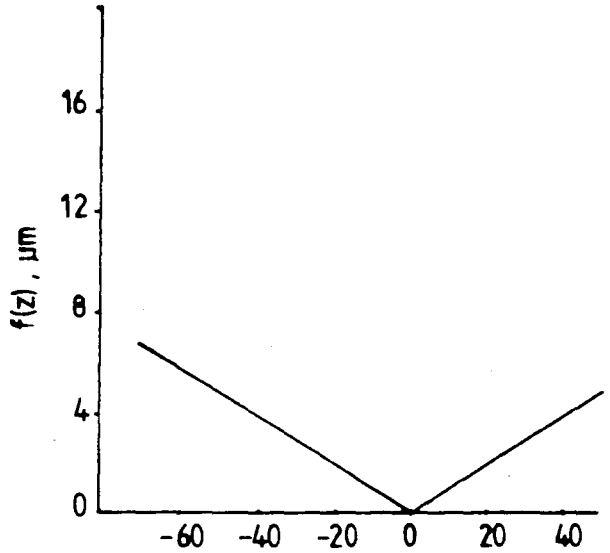
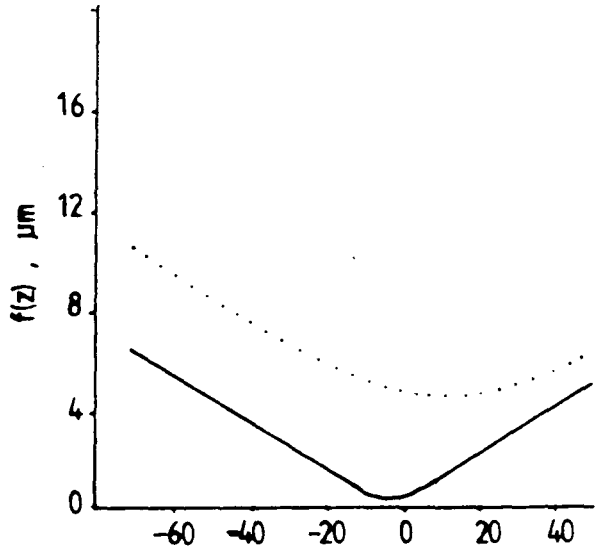


FIGURE [4.14] Longitudinal spherical aberration plot for circular-fit approximation to Sottini geodesic lens



FIGURE[4,15] L-SA PLOT FOR 20-ARC APPROXIMATE PROFILE





z = 0 corresponds to f = 18·5 mm

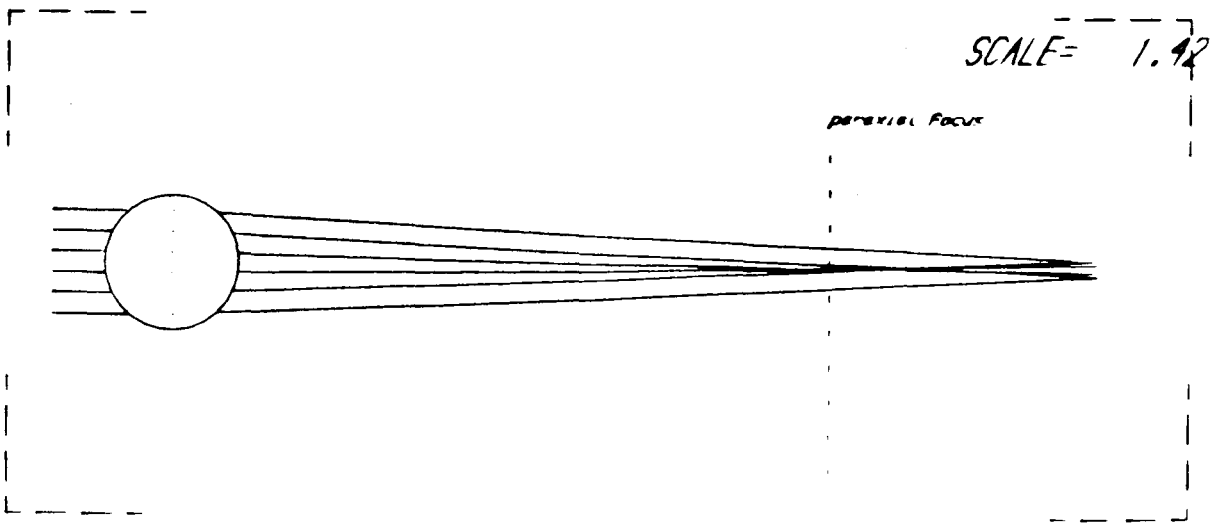
(a) (b)

FIG [4.16] The function $f(z) = \sqrt{(\sum_{i=1}^{N} y_{i,z}^2 / N)}$ where $y_{i,z}$ is the intercept height of ray i at image plane z. N is the total number of traced rays.

(a) IDEAL profile , BW = 6·0 mm

(b) PRACTICAL profile , ——— BW = 6·0 mm

..... BW = 7·2 mm



(a)

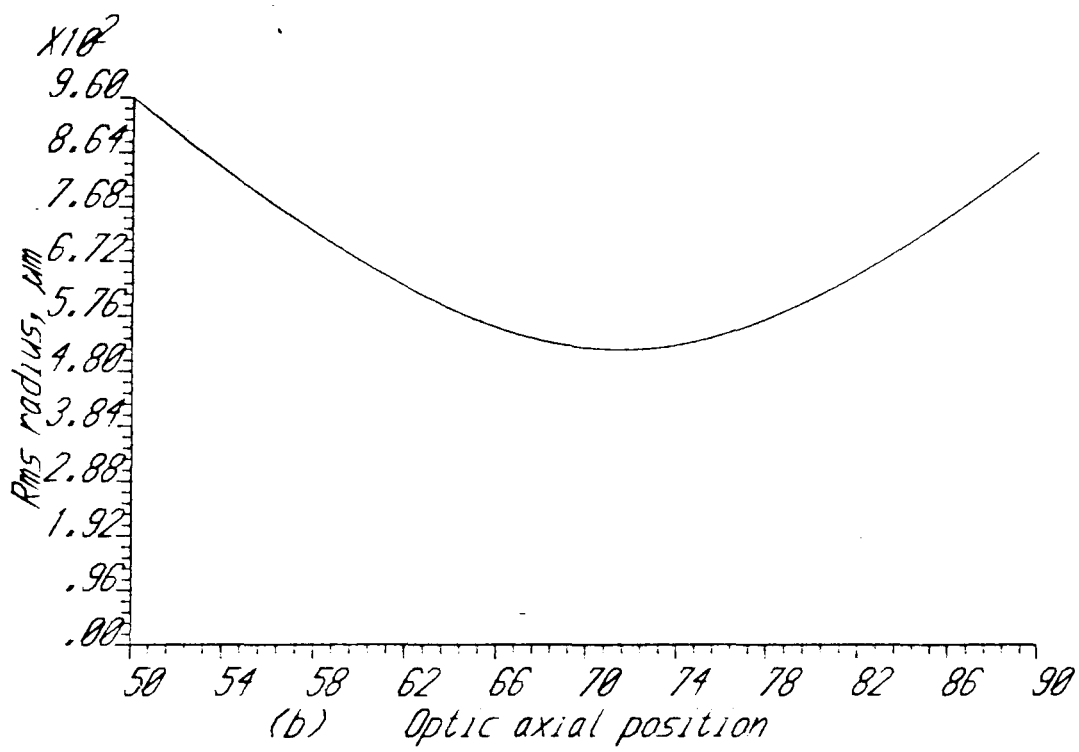


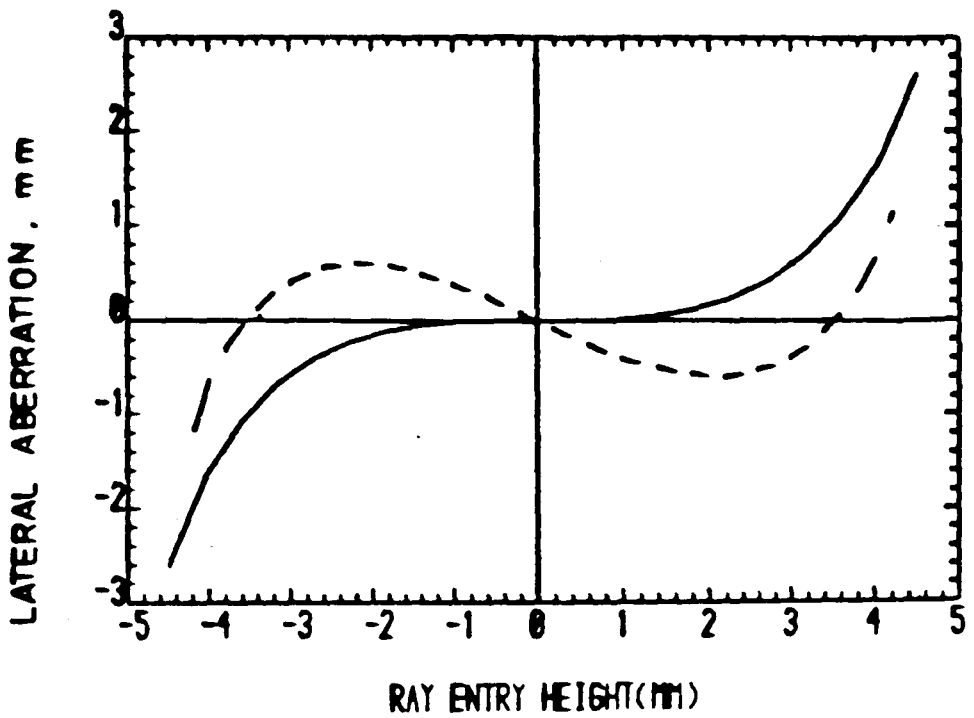
FIGURE [4,17]

(a) raytrace through integrated optical lens
lens type: SPHERICAL LENS WITH ROUNDED EDGE
lens radius= 5.23 paraxial focal length= 50.90
trace beamwidth= 8.00

(b) rms spot radius caustic

plane of least confusion= 71.41

(dimensions in mm)



——— at paraxial focus (50·9 mm)

- - - - - - at 'plane of least confusion' (71·7 mm)

FIG [4,18] Lateral aberration of lens calculated at two image planes

– overcorrected lens

focal shift required is much reduced, with the image of least confusion occurring at an axial distance of 52.2 mm, only 1.3 mm from the paraxial focus. The geometrical spot diameter is found to be just under 20 μm in such a case.

4.7 Conclusions

Perfect geodesic and Luneburg overlay lenses have been compared with possible practical realisations. In the case of overlay lenses, the practical approximation to the desired profile diplays significantly degraded optical performance, including a negative focal shift of 20%, and a large geometrical spot-size of 24 μm. A first practical approximation to the perfect geodesic lens yielded a geometrical spot-size of only 0.4 μm, but with a focal-shift of 34 μm. A second practical realisation reduced the focal-shift to 3 μm. A medium-performance long-focal-length spherical-depression lens with a toroidal rounded-edge displayed a large positive focal shift of 21.8mm at an effective f-number of f/6.36 but this was reduced to 1.3 mm at a reduced f-number of f/17, at which a geometrical blur-spot of diameter 20 μm is obtained.

REFERENCES FOR CHAPTER FOUR

[1] Stavroudis,O.N.,
'The Optics of Rays, Wavefronts and Caustics', Academic Press, New York, 1972

[2] Miyamoto,K.,
Wave optics and geometrical optics in lens design, Progress in Optics,I, 1961, 31-66

[3] Born,M. and Wolf,E.
Foundations of geometrical optics, ch III, 'Principles of Optics', Pergamon Press, Oxford, 1975

[4] Buchdahl, H.A.
'An Introduction to Hamiltonian Optics', Cambridge University Press, Cambridge, 1970

[5] Ayres,F. (ed)
'Schaum's Outline of Theory and Problems of Differential and Integral Calculus', McGraw-Hill, New York, 1972

[6] Moore, D.T.
Raytracing in gradient-index media, J. Opt. Soc. Am., 65(4), 1975, 451-455

[7] Sands,P.J.
Aberration coefficients and surfaces of best focus, J. Opt. Soc. Am., 63(5), 1973, 582-588

[8] Montagnino,L.
Raytracing in inhomogeneous media, J. Opt. Soc. Am., 58(11), 1968, 1667-1668

[9] Moore,D.T, and Stagaman,J.M.
Raytracing in anamorphic gradient-index media, Appl. Opt., 21(6), 1982, 999-1003

[10] Southwell,W.H.
Inhomogeneous optical waveguide lens analysis, J. Opt. Soc. Am., 67(8), 1977, 1004-1009

[11] Southwell,W.H.
Index profiles for generalized Luneburg lenses and their use in planar optical waveguides, J. Opt. Soc. Am., 67(8), 1010-1014 (1977)

[12] Luneburg,R.K.
'Mathematical Theory of Optics', Univ. of Calif. Press, Berkeley, California, 1964

[13] Abramowitz,M. and Stegun,I.A.
'Handbook of Mathematical Functions', Dover Publ. New York, 1965

[14] Southwell,W.H.
Geodesic optical waveguide lens analysis, J. Opt. Soc. Am., 67(10), 1977, 1293-1299

[15] Doughty, G.F.
'Aspheric Geodesic Lenses for an Integrated Optical Spectrum Analyser', Ph.D. Thesis, University of Glasgow, 1983

[16] Smith,W.J.
Image formation: geometrical and physical optics, from Section 2,'Handbook of Optics', Driscoll,W.G., Vaughan,W. (eds), McGraw-Hill, New York, 1978

[17] Nijboer,B.R.A.
The diffraction theory of optical aberrations. Part I: General discussion of the geometrical aberrations, Physica, X(8), 1943, 679-692

[18] Vahey,D.W.
'Optical Waveguide Geodesic Lenses', Final Report, Air Force Avionics Lab., Ohio,USA, AFAL-TR-78-85, 1978

CHAPTER FIVE

THE BEAM PROPAGATION METHOD

CHAPTER FIVE COMPUTATIONAL ANALYSIS:II

THE BEAM-PROPAGATION METHOD (BPM)

5.1 Introduction

A geometrical optics approach to the propagation of light in lenses, and in the homogeneous space surrounding the lenses, reveals much useful information about the lens properties, as was found in chapter two and chapter four. In particular, the a priori design of a lens can only be carried out using the principles of geometrical optics. However, in physical optics, phenomena associated with diffraction exist which are not adequately described by geometrical optics, and for which a second-order solution of wave propagation is required. Diffraction calculations using classical methods are, generally, very complicated, even in homogeneous media. Diffraction effects associated with propagation in inhomogeneous media are rarely considered.

A simple algorithm for calculating the propagation of wave-fields in inhomogeneous media has been developed over recent years which relies upon numerical methods for solution. The algorithm is known as the Beam Propagation Method, or BPM. The BPM is based on solutions of the scalar wave-equation in the small-index-variation and paraxial approximations. It uses the elegant and extremely powerful techniques of Fourier optics extensively. An important feature of the method is that no special assumptions need be made about the form of the incident fields. Furthermore, large wave aberrations could be modelled, conceivably.

In this chapter, the Fourier optics representation of wave-fields will be outlined, together with methods of implementing the Fourier representation on a computer. A well-known algorithm for the propagation of wave-fields in homogeneous media will then be given. Propagation in inhomogeneous space, such as is found in the lenses considered in this thesis, will then be considered and the BPM will be introduced. A single lens design will then be modelled using the BPM, under a variety of incident field conditions. Previous published work on the use of the BPM in relation to inhomogeneous lenses has not dealt with the detailed evolution of the field in the region of the focus. If the BPM is to be of use in the investigation of inhomogeneous lenses, it must adequately represent the focal field. Field patterns obtained using the present BPM model indicate that a negative focal shift is introduced which is compatible with recent results published by other workers.

5.2 The angular spectrum of plane waves

The following discussion is similar to those found in the texts of Goodman [1] and Gaskell [2]. Consider a wave-field propagating in a medium which extends over three-dimensional space. If the medium is homogeneous, so that its properties do not vary from point-to-point, and isotropic, so that its properties do not vary with direction, the propagation of the wave-field in the medium may be described very easily using Fourier transform techniques.

The wave-field is assumed to be monochromatic and linearly-polarized. The time variation of the field may be neglected due to its periodicity and the wave-field conveniently represented by its complex amplitude only. The complex amplitude describes the spatial variation of the magnitude and phase of the field and is given by the scalar function of position:

$$u(x,y,z) = a(x,y,z)\ \exp(j\,\phi(x,y,z)) \qquad (5.1)$$

As usual, the real part of equation (5.1) represents the real physical wave.

A wave-field at a constant plane $z=z_i$ may be represented by a function $u_i(x,y)$. If the wave-field is physically realizable, it will possess a two-dimensional Fourier transform at this plane:

$$U_i(\xi,\eta) = FF(u_i(x,y)) \tag{5.2}$$

Consequently $u_i(x,y)$ may be expressed as the inverse Fourier transform of $U_i(\xi,\eta)$:

$$u_i(x,y) = \int_{-\infty}^{\infty}\int U_i(\xi,\eta)e^{j2\pi(\xi x + \eta y)}d\xi d\eta \tag{5.3}$$

Now a plane wave propagating with direction cosines (α,β,γ) may be represented by a function of the form:

$$\text{Plane wave} = Ae^{jk(\alpha x + \beta y)}$$

$$= Ae^{j2\pi(\alpha x/\lambda + \beta y/\lambda)} \tag{5.4}$$

where A is a constant indicating the peak amplitude of the wave. If $\alpha = \lambda\xi$ and $\beta = \lambda\eta$, the exponential term in equation (5.2) may be regarded as a unit-amplitude plane wave propagating with direction cosines $(\lambda\xi, \lambda\eta, \sqrt{1 - \lambda^2(\xi^2 + \eta^2)})$. In accordance with the usual understanding of the Fourier transform , $u_i(x,y)$ may then be regarded as a linear superposition of plane wave components travelling in directions governed by the values of the direction cosines. The amplitude of each plane wave component is governed by the weighting function $U_i(\xi,\eta)$. Thus $U_i(\xi,\eta)$ is referred to as the angular spectrum of $u_i(x,y)$ and is given by:

$$U_i(\xi,\eta) = \int_{-\infty}^{\infty}\int u_i(x,y)e^{-j2\pi(\xi x + \eta y)}dxdy \tag{5.5}$$

Each plane wave component is infinite in extent and propagates

with a constant amplitude, but with varying phase. Since the direction of propagation for each plane wave component is different, each component undergoes a different phase change as the field propagates from one plane for which z = constant to another. At some plane $z_1 > z_i$, the phase will have increased by an amount:

$$k\gamma(z_1-z_i) = k(z_1-z_i)\sqrt{[1 - \lambda^2(\xi^2+\eta^2)]} \qquad (5.6)$$

Since the amplitude remains unchanged, a transfer function can be easily defined which specifies the propagation of a wave-field from the plane $z=z_i$ to the plane $z=z_1$:

$$\frac{U_1(\xi,\eta)}{U_i(\xi,\eta)} = \exp\{jk(z_1-z_i) \quad [1 - \lambda^2(\xi^2+\eta^2)] \} \qquad (5.7)$$

The benefits of representing an optical wave-field in terms of its angular spectrum are those classically accruing from the Fourier transform, ie the individual components are treated much more simply than the whole and the reconstruction of the field is simply carried out using the principle of superposition.

5.3 Computational representation of a complex 1-D wave-field in spatial and angular frequency domain

In integrated optical lenses, the field variation in the dimension perpendicular to the plane of the wave-guide is usually considered unimportant, so that the propagation of a one-dimensional wavefield u(x) as a function of a variable z may be considered.

In order to perform computations on such wave-fields it is necessary to have a representation of the field consisting of discrete sampled values. The angular spectrum must also be sampled in this way. The continuous Fourier transform and its inverse may then be represented by the discrete Fourier transform (DFT) pair [3].

The sampling interval required to avoid aliasing in both domains is given by the well-known Shannon criterion, which states that the interval must be less than or equal to half of the smallest period present in the signal, for the sampled signal to faithfully reproduce the original. 'Interval' and 'period' are used here instead of the more usual term 'frequency', to avoid possible ambiguities in the interpretation of sampling in both spatial and angular domains.

Optical signals are usually of limited extent in the transverse direction in space due to the presence of a limiting aperture. If a signal is band-limited in one domain, it cannot be band limited in the other [3], so that a 'smallest period' cannot be defined. In practice, a period can be defined below which a negligible amount of signal energy is contained. The signal can then be sampled with arbitrarily little distortion being introduced.

Sampling of the truncated field introduces periodicity in the spectral domain, and vice versa. The wave field and its spectrum then have to be truncated numerically so that only a finite number of samples N, extending over one period of the sampled signal is contended with. Sampling and truncation together introduce distortion in the signal. This is minimised by choosing the largest truncation interval and smallest sampling interval possible, commensurate with the speed and accuracy required of the calculations.

The DFT pair is therefore an approximation to the continuous Fourier transform pair:

$$u_i(x) = \int_{-\infty}^{\infty} U_i(\xi) e^{j2\pi\xi x} d\xi$$

$$U_i(\xi) = \int_{-\infty}^{\infty} u_i(x) e^{-j2\pi\xi x} dx. \qquad (5.8)$$

The angular spatial frequency, or wavenumber, k_x is more commonly used than the spatial frequency variable ξ . The relationship

between them is:

$$k_x = 2\pi\xi \tag{5.9}$$

The sampled field and spectrum both consist of N equidistant samples specified for computational purposes by:

$$u_i(m\Delta x) = \frac{1}{N}\sum_{n=0}^{N-1} U_i\left(\frac{n}{N\Delta x}\right)e^{j2\pi nm/N}$$

$$m = 0,1,\ldots,N-1$$

$$U_i(n/N\Delta x) = \sum_{n=0}^{N-1} u_i(m\Delta x)e^{-j2\pi nm/N}$$

$$n = 0,1,\ldots,N-1 \tag{5.10}$$

The width of the field in the spatial domain is $(N-1)\Delta x$, where Δx is the sampling interval in this domain. The corresponding width of the field in the spectral domain is $(N-1)/(N\Delta x)$.

Inspection of (5.10) reveals that each of the N samples in the transform require N multiplications and N additions, so that $2N^2$ arithmetic operations are required to calculate the complete spectrum. However, the Fast Fourier Transform algorithm developed by Cooley and Tukey [3], which was originated in work by Gauss [4], calculates the DFT with accuracy and greatly increased efficiency, since only NlogN operations are required owing to a clever matrix decomposition. The FFT allows numerical algorithms for modelling the propagation of optical wave-fields to be implemented with relative ease.

5.4 An algorithm for computing the propagation of optical wave-fields in homogeneous, isotropic space

The previous two sections, on the angular spectrum of plane

waves and on computational implementations of the Fourier transform may be taken together to model the propagation of arbitrary optical wave-fields in homogeneous, isotropic media. A simple algorithm might be:

1. Sample the optical wave-field at the input-plane of interest and obtain the angular spectrum using the forward FFT.

2. Propagate each plane-wave component separately to the next plane of interest, using the transfer function given in equation (5.7). The angular spectrum at the new plane is then obtained.

3. Construct the optical wave-field at the new plane by taking the inverse FFT.

Such an algorithm has been found useful in the solution of many optical problems. In particular, it has proved competitive with evaluations of the Kirchhoff diffraction integral in near- and far-field diffraction problems, as shown by Sziklas and Siegman [5].

5.5 Theory of the beam-propagation method

The algorithm given above for the propagation of waves in homogeneous, isotropic space cannot be used directly for the modelling of propagation in inhomogeneous space, since plane waves do not remain plane in such a case. It is desirable therefore to obtain a new algorithm for describing propagation in inhomogeneous space. The beam propagation method (BPM) is just such an algorithm. It is essentially a numerical method for solving the scalar wave equation in inhomogeneous media.

For the purposes of the present work, the main interest in the use of the BPM lies in its ability to describe the development of the optical field in all its complexity within

inhomogeneous lenses without resorting to the methods of geometrical optics. Particularily, the effects of diffraction can be modelled very accurately without, importantly, making any assumptions about the field at the exit pupil of the lens. The usual methods for investigating diffraction effects generally display three important characteristics

(i) the observation distances are large, and the angles with respect to the optical axis are small;

(ii) the lens is considered to be thin, so that the amplitude distribution at the exit pupil can be assumed to be either constant or Gaussian, with the effects of beam-truncation also included;

(iii) aberration effects are rarely modelled so that the phase-front leaving the lens is assumed to be perfectly spherical.

The BPM displays, to some extent, the first characteristic but not the other two. The fact that the amplitude variation can be arbitrarily specified at the entrance pupil of a thick lens is important as it allows apodization effects to be modelled.

The problem is to calculate the propagation of a wave-field $u(x,z)$ through a medium with a refractive index function $n(x,z)$, given an initial wave-field $u_i(x,z)$. Van Roey et al [6] have derived a general theory of the BPM using a Green's function approach. The method is complicated, and a mathematical text such as that of Arfken [7] is indispensable as an aid to understanding it. A simpler derivation is possible however, and the published theories of Feit and Fleck [8,9] and Lagasse [10] are used in the following discussion. The theory given by Feit and Fleck is attractive in that the basic features of the technique are clearly laid out, for a 3-D medium with a refractive index variation in the directions transverse to propagation only. Lagasse extends the theory to media with a refractive index variation in the direction of propagation. The following theory

will be restricted to two-dimensional spaces.

Three important assumptions are made:

(i) the problem has to be described adequately by the scalar wave equation, thus restricting the discussion to situations where separable (uncoupled) solutions can be derived for the TE and TM modes;

(ii) the refractive index can be written as a sum of unperturbed and perturbed parts:

$$n(x,z) = n_o(x) + \Delta n(x,z) \qquad (5.11)$$

where $\Delta n << n_o$. A purely transverse index variation means that the perturbation is , further, a function of the variable x only. $n_o(x)$ must be chosen such that the solutions of the scalar wave equation:

$$\nabla^2\phi + k^2 n_o^2(x)\phi = 0 \qquad (5.12)$$

are known eigenfunctions, $\phi_n(x)e^{-jk_n z}$. If $n_o(x)$ is chosen to be constant the eigenfunctions are given by the angular spectrum of plane waves.

(iii) the variation of n(x,z) along z, if there is to be any, must not contain any sharp discontinuities or periodicities, both of which give rise to reflected waves which are not allowed with the BPM.

Consider first a medium in which the index variation in the z-direction is constant. The scalar (Helmholtz) wave equation in such a medium is:

$$\nabla^2 u + k_o^2 n^2(x) u = 0 \qquad (5.13)$$

where $\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial z^2}$, and k_o is the free-space wave-number.

The standard procedure for dealing with this equation is to assume:

$$u(x,z) = G(x,z)e^{-jk_o nz} \qquad (5.14)$$

with $G(x,z)$ a slowly varying function. Substituting (5.14) into (5.13) and neglecting second-order terms in z, one recovers a first-order differential equation in z which is variously known as the paraxial, parabolic or Fresnel form of the wave equation. This approach is used by Kogelnik and Li, for example, in their study of the propagation of Gaussian beams [11].

Instead of going for the Fresnel approximation directly, consider that the solution at $z = \Delta z$ may be written formally in terms of the field at z=0 as:

$$u(x, z) = \exp[+j\Delta z(\nabla_\perp^2 + k_o^2 n^2)^{1/2}]u(x,0) \qquad (5.15)$$

where: $\nabla_\perp^2 = \dfrac{\partial^2}{\partial x^2}$

(Equation (5.15) may easily be shown to lead to the wave equation (5.13) if both sides of (5.15) are partially differentiated with respect to z). The square root in (5.15) can be rewritten:

$$(\nabla_\perp^2 + k_o^2 n^2)^{1/2} = \frac{\nabla_\perp^2}{(\nabla_\perp^2 + k_o^2 n^2)^{1/2} + k_o n} + k_o n \qquad (5.16)$$

If n in the first right-hand member of (5.16) is replaced by n_o, the unperturbed index, (5.16) becomes:

$$(\nabla_\perp^2 + k_o^2 n^2)^{1/2} \simeq \frac{\nabla^2}{(\nabla_\perp^2 + k^2)^{1/2} + k} + k + k[(n/n_o)-1] \qquad (5.17)$$

where $k = k_o n_o = n_o \omega/c$, with ω being the angular frequency and c being the velocity of light.

The approximation made in (5.17) is valid for sufficiently small perturbations in $n(x,z)$.

It is now convenient to express $u(x,z)$ as the product of a complex field amplitude $v(x,z)$ and a carrier wave moving in the positive z-direction:

$$u(x,z) = v(x,z)e^{-jkz} \tag{5.18}$$

Substitution of (5.18) into (5.15) and taking the negative sign, indicating forward propagation, gives:

$$v(x,\Delta z) = \exp\left\{-j\Delta z\left[\frac{\nabla_\perp^2}{(\nabla_\perp^2+k^2)^{1/2}+k} + \chi(x)\right]\right\}v(x,0) \tag{5.19}$$

where $\chi(x) = k[(n(x)/n_o)-1]$

(5.19) can be rewritten in symmetrized split operator form, to second order in Δz, as:

$$\begin{aligned} v(x,\Delta z) = \exp&\left\{\frac{-j\Delta z}{2}\left[\frac{\nabla_\perp^2}{(\nabla_\perp^2+k^2)^{1/2}+k}\right]\right\}\exp(-j\Delta z\,\chi) \\ &\times \exp\left\{\frac{-j\Delta z}{2}\left[\frac{\nabla_\perp^2}{(\nabla_\perp^2+k^2)^{1/2}+k}\right]\right\}v(x,0) \\ &+ O\ (\Delta z)^3 \end{aligned} \tag{5.20}$$

The splitting of the operators results in a separation of the propagation part of the calculation from another part which may be recognised as a correction factor to allow for the perturbation of the index. The operators inside the brackets of (5.19) do not commute, and so there is an error term since an approximation is invoked that holds only for limited propagation distances Δz. The operator splitting and its consequences are discussed in detail in reference [9].

The operation:

$$\exp\left\{-j\Delta z\left[\frac{\nabla_\perp^2}{(\nabla_\perp^2+k^2)^{1/2}+k}\right]\right\}v(x,\ 0)$$

is equivalent to solving the Helmholtz wave equation for:

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial z^2} + k_o^2 n_o^2\, u = 0 \qquad (5.21)$$

with $u(x,0)$ as an initial condition, so that the operator defines the propagation in the unperturbed medium. If $\nabla_\perp^2$ is neglected in comparison with $k_o^2 n_o^2$ in the denominator of (5.19) and (5.20), one recovers the paraxial, or Fresnel approximation:

$$v(x,\Delta z) = \exp\left\{-j\Delta z\left[\frac{\nabla_\perp^2}{2k} + \chi(x)\right]\right\}v(x,0) \qquad (5.22)$$

The propagation in the unperturbed medium may be calculated exactly for the case of n_o = constant by finding the angular spectrum of plane waves and propagating the plane waves using the transfer function (5.7) . Advancing the solution for $v(x,z)$ by repeated application of (5.21) is equivalent to propagating the beam through a periodic array of thin lenses, as shown in Figure [5.1] . The first lens is located at $z = \Delta z/2$ and the remaining lenses are separated from one another by Δz. Each lens imposes a phase-front:

$$\begin{aligned}\phi(x) &= \Delta z\ \chi(x)\\ &= \Delta z k[(n/n_o)-1]\\ &= \Delta z k_o[n-n_o]\\ &= \Delta z k_o \Delta n \qquad (5.23)\end{aligned}$$

on the beam and the propagation between the lenses is given by (5.21).

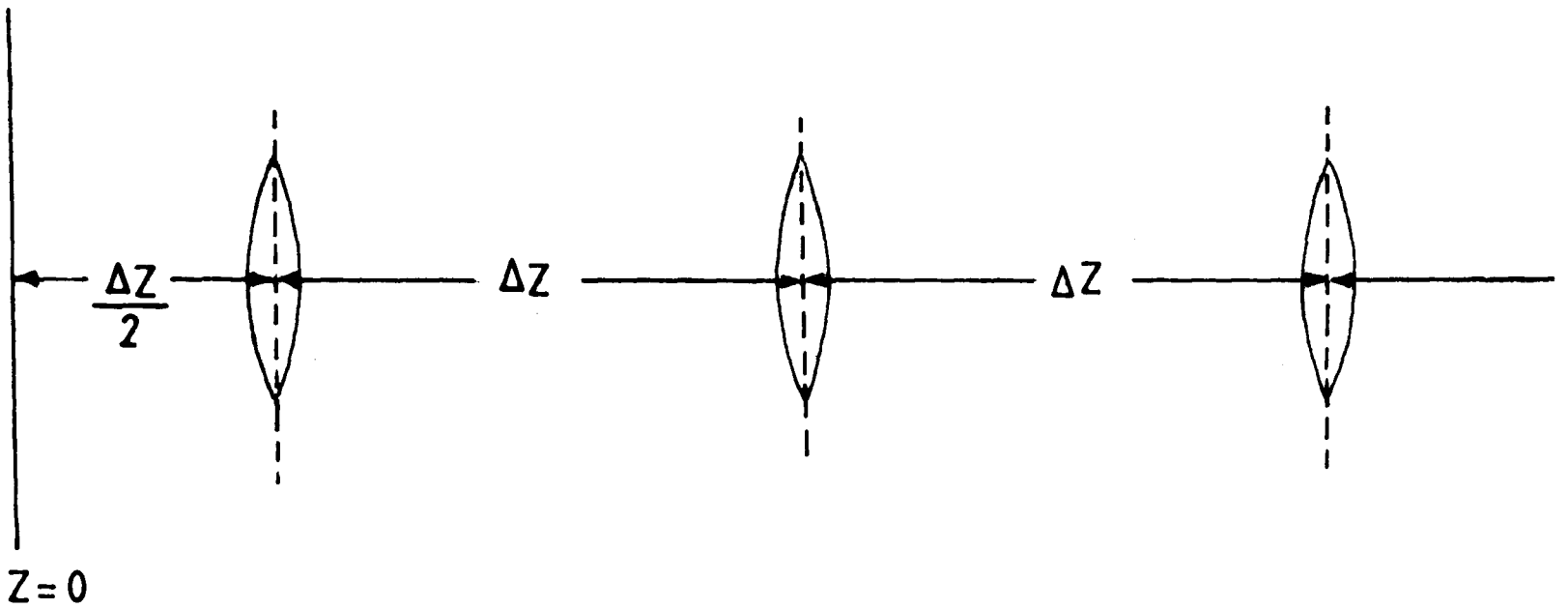


FIGURE [5,1] BPM algorithm replaces variable-index medium by periodic array of thin-lenses. In between lenses field satisfies the Helmholtz equation for a homogeneous medium.

It may be recalled that the above analysis was derived for the case of a medium with a refractive index variation in the transverse direction only. The major advantage of studying this case is that the underlying reasons for treating an inhomogeneous medium with a small and slow variation in refractive index as a periodic array of thin lenses are clearly illustrated. Lagasse [10] goes on to treat the case of propagation in a 2-D medium with a refractive index variation in both x and z directions. Let it be assumed that the principle of propagation in the unperturbed medium followed by the application of a correction factor to allow for the index variation is still valid.

The field $u(x,\Delta z)$ is written, therefore, as the product of the field propagating in the unperturbed medium n_o and a phase-updating correction factor e^{Γ} :

$$u(x,z_o+\Delta z) = u'(x,z_o+\Delta z)e^{\Gamma} \qquad (5.24)$$

where u' satisfies:

$$\nabla^2 u' + k_o^2 n_o^2 u' = 0 \qquad (5.25)$$

with:

$$u'(x,z_o) = u(x,z_o) \qquad (5.26)$$

and the solutions of (5.25) are the eigenfunctions:

$$u'(x,z) = \sum_{n=1}^{\infty} B_n^{\,o}\, \phi_n(x) e^{-jk\Delta z} \qquad (5.27)$$

with:
$$B_n^{\,o} = \int_{-\infty}^{\infty} u(x,z_o)\, \phi_n^{\,*}(x)dx \qquad (5.28)$$

In the case n_o = constant, equations (5.27) and (5.28) represent Fourier transforms. The scalar wave equation is:

$$\nabla^2 u + k_o^2 n^2(x,z) u = 0 \qquad (5.29)$$

Substituting (5.24) into (5.29), and taking (5.25) into

account, yields:

$$u'\nabla^2\Gamma + u'|\nabla\Gamma|^2 + 2\nabla u\nabla\Gamma + k_o^2(n^2-n_o^2)u' = 0 \quad (5.30)$$

Since Δz can be chosen small, a series expansion for Γ is possible:

$$\Gamma(x,z) = \sum_{n=1}^{\infty} A_n(x)\ z^n \quad (5.31)$$

Substitution of (5.31) into (5.30) and equating the coefficients of z^n to zero yields a set of equations for the A_n. The first equation, from the constant term z^0, is:

$$A_1^2 + \frac{2}{u'}\frac{\partial u' A_1}{\partial z} + k_o^2(n^2-n_o^2) + 2A_2 = 0 \quad (5.32)$$

The form of A_1 needs to be determined. For this some simplifying approximations are required. The presence of A_2 complicates matters, and it would be beneficial if its effect could be ignored. Assume therefore that:

$$A_2 << \frac{k_o^2(n^2-n_o^2)}{2}$$

This condition can be justified by calculating A_2 from the second equation resulting from the series expansion. (5.32) is then reduced to a quadratic equation for which the solution is:

$$A_1 = \frac{1}{u'}\frac{\partial u'}{\partial z}\left[-1 + \left[1-\left(\frac{1}{u'}\frac{\partial u'}{\partial z}\right)^{-2}k_o^2(n^2-n_o^2)\right]^{1/2}\right] \quad (5.33)$$

It is possible to calculate $\partial u'/\partial z$ in the spectral domain, if necessary. However, if

$$k_o^2\, n^2\left(\frac{1}{u'}\frac{\partial u'}{\partial z}\right)^{-2} << 1$$

is assumed, then:

$$A_1 = \frac{-k_o^2 (n^2-n_o^2) u'}{2 \quad u'/z} \qquad (5.34)$$

Introducing paraxiality enables one to assume that u' consists of a wave travelling along the z-axis:

$$u' = e^{-jk_o n_o z}$$

Coefficient A_1 then becomes:

$$A_1 = \frac{-jk_o}{2n_o} (n^2-n_o^2)$$

$$= \frac{-jk_o (n-n_o)(n+n_o)}{2n_o}$$

$$= -jk_o \Delta n, \quad \text{if } (n+n_o) = 2n_o. \qquad (5.35)$$

As in Feit and Fleck's theory the last approximation is valid for sufficiently small index perturbations.

The correction factor Γ is finally given by:

$$= -jk_o \Delta n \Delta z \qquad (5.36)$$

which is of the same form as that given for media with only a transverse variation in refractive index in equation (5.23) and is equivalent to the effect of a thin lens.

The above developments are not unexpected, as the representation of an inhomogeneous medium as a periodic array of thin lenses within the approximations of paraxialty and small index variation is well known (see, for example, reference [11]).

Classical theories of diffraction have problems in dealing with three types of complication:

(i) complications in the field incident on the diffracting aperture. Gaussian, truncated-Gaussian and uniform amplitude distributions are generally the only cases considered;

(ii) complications in the phase change introduced by the lens. Lenses almost always have to be thin. The thin lenses are generally assumed to give rise either to perfect phase changes or to small phase distortions described by simple analytic functions. Inhomogeneous lenses are thick, however, and can give rise to complicated phase distortions. For example, fabrication processes could perturb the refractive index functions in complex ways;

(iii) complications in the amplitude transmission-function of the lens Lenses are generally not considered to introduce any amplitude distortion, or are assumed to introduce very simple one-dimensional variations of amplitude.

The great advantage conferred by the BPM is that all of the above problems can be dealt with simply, in principle, so long as the conditions of validity are satisfied.

5.6 Conditions for the applicability of the BPM

Van Roey [6] has listed several conditions for the applicability of the BPM. The conditions are:

$$\Delta n_{max} \sin^2 \alpha \, \frac{\Delta z}{\lambda} \ll 1 \, , \tag{5.37}$$

where Δn_{max} is the maximum perturbation in the refractive index over the propagation interval Δz, α is the maximum angle corresponding to a significant part of the angular spectrum of the forward-propagating beam and λ is the free-space wavelength;

$$p \gg \frac{\lambda \sqrt{2}}{n_o} \, , \tag{5.38}$$

where p corresponds to the period of the highest spatial frequency component of the refractive index profile $\Delta n^2(x)$, and n_o is the unperturbed reference refractive index over the propagation interval;

$$p >> \frac{1}{n_o} (2\pi \lambda |\Delta n|_{max} \Delta z)^{1/2} , \qquad (5.39)$$

$$\frac{\pi \tan\alpha \, \Delta z}{p} << 1 , \qquad (5.40)$$

$$\left(\frac{2\pi \tan\alpha}{p\lambda} |\Delta n|_{max}\right)^{1/2} \Delta z << 1 \qquad (5.41)$$

These conditions amount to restrictions on the maximum index perturbation, and its gradient, and the maximum spatial frequency which the optical wave-field is allowed to take. The conditions can be used to calculate a propagation step-size suitable for propagating a wave-field through a particular structure, given the refractive index profile of the structure and the spatial frequency profile of the incident optical field.

The examples of interest are inhomogeneous waveguide lenses which can have fairly large index perturbations ($\Delta n = 0.1 n_o$) which vary slowly with wavelength, however. It turns out that the first condition, (5.37), is the most stringent in such a case. If the operating wavelength in vacuum is 633 nm, $\Delta n = 0.22$ and the maximum angular aperture is 15 degrees, corresponding to the paraxiality limits, condition (5.37) gives:

$$\Delta z << 43 \, \mu m$$

If $\Delta z = 10$ μm is taken as satisfying the condition, a BPM model of an inhomogeneous lens of 10.0 mm diameter then requires 1000 thin-lens elements and 1000 homogeneous spaces for accurate modelling.

5.7 The sampling interval

The decomposition of the wave-field into the angular spectrum of plane waves is computed using the Discrete Fourier Transform (DFT). In the inhomogeneous region the DFT is implemented using the Fast Fourier Transform, for reasons of computational efficiency. Two constraints on the field sampling interval, Δx, exist:

(i) the Sampling Theorem states that the spatial sampling rate must be greater than, or equal to, twice the highest spatial frequency present in the angular spectrum, for the sampled field to faithfully reproduce the properties of the original field. It can be shown that an optical system with focal length f, truncated by an aperture of radius, a, and operating at a free-space wavelength λ in a medium with refractive index n can only transmit spatial frequencies ξ such that:

$$\xi \leqslant n\, a/f\lambda \qquad (5.42)$$

[12] . Equation (5.42) sets a maximum value of spatial frequency for the optical field and, consequently, a maximum value for the sampling interval. In reality, it is advisable to make the sampling interval smaller than the maximum value, so as to introduce a guard-band in the spectral domain which helps to prevent aliasing.

(ii) the spatial sampling interval must not be so small, however, that the condition of paraxiality no longer holds. Paraxiality, rather conservatively perhaps, may be assumed to hold for plane-wave components of the spectrum that do not deviate by more than 10 degrees from the optical axis. This sets a minimum limit of approximately f/2.85 on the f-number that can be modelled using the BPM. In terms of the corresponding maximum spatial frequency:

$$\xi_{max} = \frac{1}{2\,\Delta x_{min}} = \frac{n \quad \tan 10^{o}}{\lambda\,(1 + \tan^{2}10^{o})^{1/2}} \qquad (5.43)$$

Conditions (i) and (ii) may be taken together as follows:

$$\frac{\lambda(1 + \tan^2 10^o)^{1/2}}{2n \tan 10^o} < \Delta x < \frac{f\lambda}{2na} \qquad (5.44)$$

As an example, consider n = 2.206, λ = 0.633 μm, f = 18.5 mm and a = 1.5 mm. Then:

$$0.83\ \mu m < \Delta x < 1.77\ \mu m$$

These limiting conditions on the sampling interval are shown in schematic form in Figure [5.2].

5.8 Implementation of the BPM for inhomogeneous lens structures

The BPM is implemented in inhomogeneous lens structures in four stages.

1. An incident field is created, in complex amplitude form. The field can be adjusted to any required incident angle or offset distance from the axis. The most commonly used field is of Gaussian shape, although other field shapes can be used. The Gaussian, having a beam diameter 2w, is truncated by a "hard-aperture" of diameter 2a with transmission properties

$$t = 1, \quad |x| \leqslant a$$
$$t = 0, \quad |x| > a$$

The consequences of using such an aperture in conjunction with the DFT have been discussed in reference [9]. Generally, the dimensions of the aperture are chosen so that a large guard-band with zero-field exists in the vicinity of the boundaries of the computational grid. For example, a 4096 x 1.5 um grid is used to initiate propagation through a lens of a plane wave-field truncated by an aperture of width 3.0mm. Thus only 2000 grid points are contained within the aperture, with the remaining 2096 serving as a guard-band. A similar guard-band is present in the

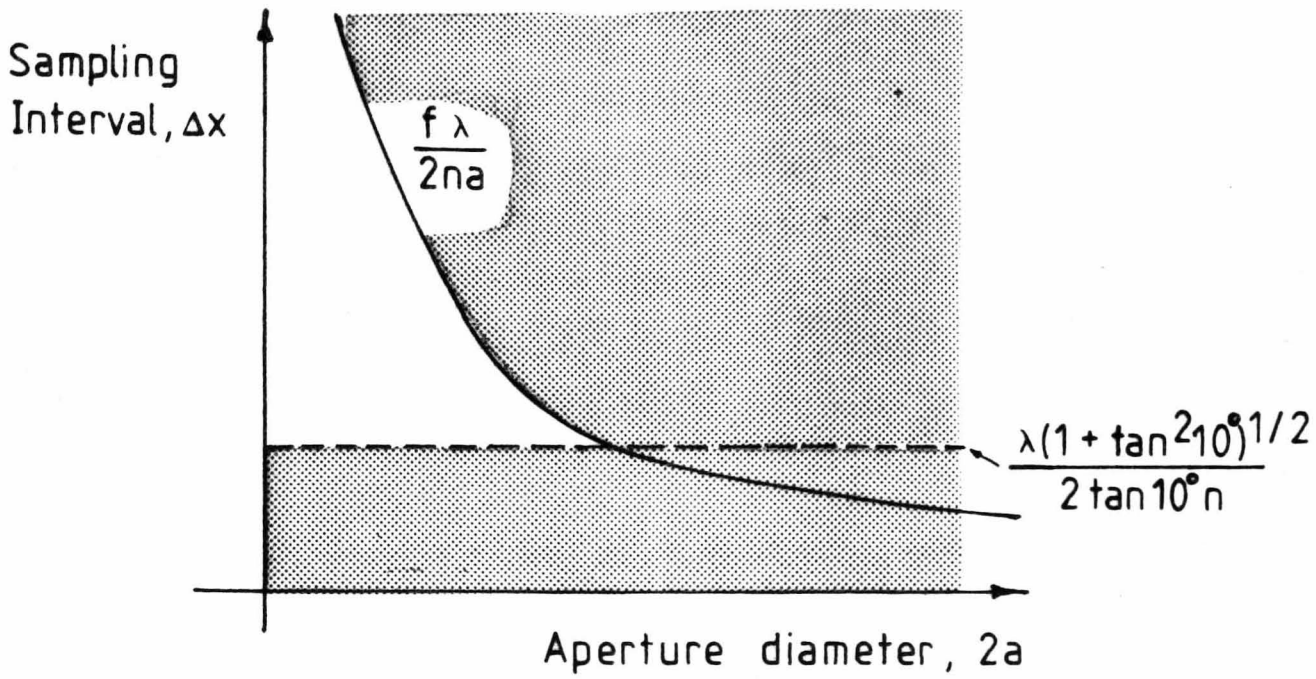


FIGURE [5, 2] LIMITS ON BPM SAMPLING INTERVAL

angular spectrum. Such guard-bands help to ensure accuracy with the DFT.

2. The chosen lens-structure, whether it be of the overlay type or the geodesic type, is transformed into an equivalent variable refractive index lens. The profile is specified at a large number of points, with linear interpolation being used to calculate the refractive index between these points.

3. The field is propagated through the lens, using the alternating propagation/phase-update components of the BPM algorithm. The field is monitored in both configurational and spectral spaces to ensure that significant amounts of energy within the band-limited waveforms do not approach the grid-boundaries, a condition which effectively results in an aborted solution [9]. A truncated plane wave-field propagating through a geodesic lens, computed using the BPM, and showing the extent of the guard-bands is given in Figure [5.3]. Because of the guard-bands, neither the field nor the spectrum approach the computational grid-boundaries.

4. Propagation in the homogeneous region beyond the lens is carried out using the algorithm given in section (5.4). The DFT is used directly in this part of the calculation. A great deal of unnecessary information about transverse points well away from the focus would be calculated if the FFT were used, since the FFT requires that the computational grid be of fixed width throughout the calculation. No such requirement exists with the direct use of the DFT. Furthermore, the spacing between the grid-points must remain fixed at the spacing set at the lens input with the FFT. Again, no such requirement exists with the DFT, so that a great deal of detailed information about the field in the focal region can be obtained.

5.9 Results

Van der Donk et al [13,14] have previously used the BPM to

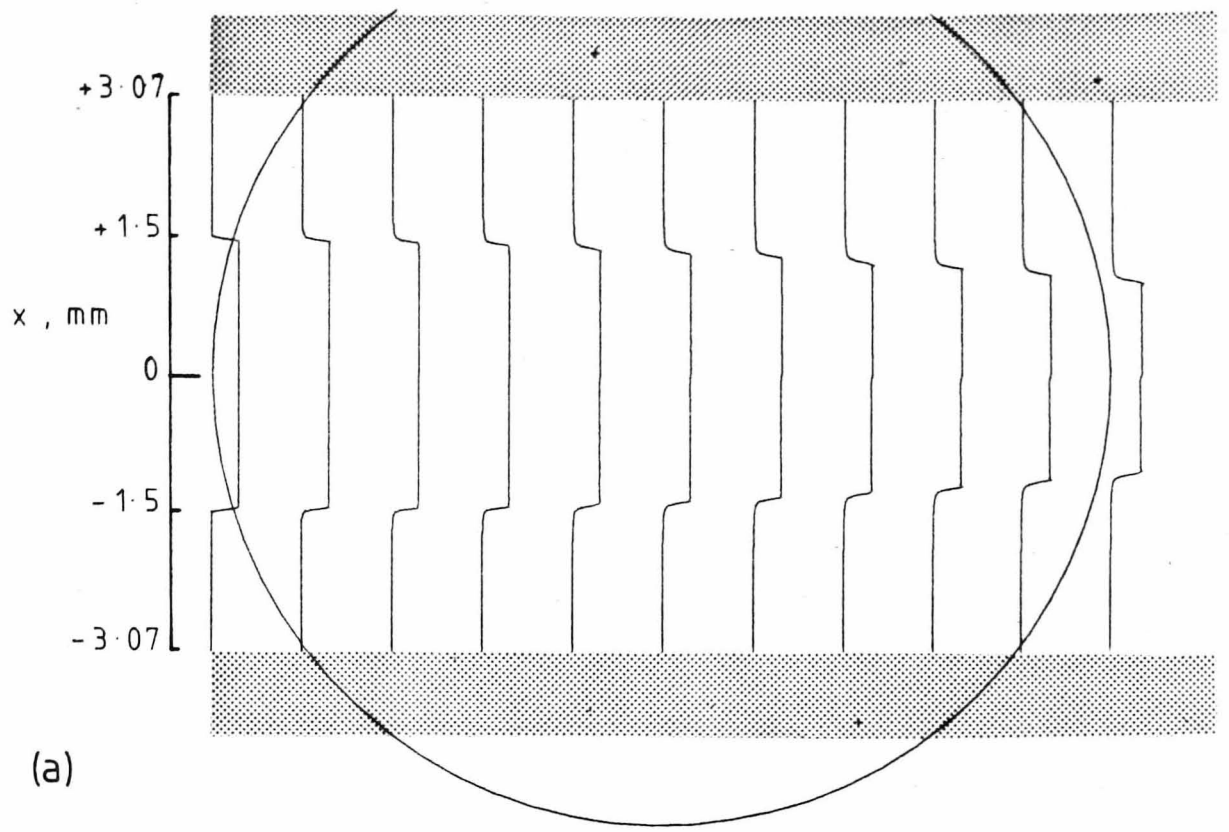


FIGURE [5, 3] Beam propagating through geodesic lens with computational grid boundaries shown (a) spatial field (b) angular spectrum

model the propagation of wave-fields through geodesic lenses. They did not report, however, on the detailed structure of the field in the image space, particularly in the focal region. The correct prediction of the focal field, particularly in the diffraction-limited case, is an absolute necessity if the BPM is to be of real value in the analysis of inhomogeneous lenses.

The aspheric geodesic lens specified in chapter two, having theoretically perfect focusing properties, was chosen as a demanding test for the BPM. The lens had an overall diameter of 10.0 mm, an inner region of 7.4 mm diameter which constituted the useful part of the lens, and a focal length of 18.5 mm. The lens had a relatively small effective f-number , therefore, of f/3. Since geodesic lenses are free of chromatic aberration, the choice of operating wavelength is arbitrary. A wavelength of 0.633 μm was chosen for the BPM tests since optical experiments would be carried out at this wavelength. The refractive index of the homogeneous medium outside the lens was chosen to be 2.2065, a value comparable with the effective refractive index of single-mode titanium-diffused guides in $LiNbO_3$. The lens and the surrounding region are further assumed to be isotropic; $LiNbO_3$ is anisotropic, but the inclusion of such a feature would have complicated the investigation unnecessarily, since it was the BPM itself that was under test rather than a given lens. Van der Donk has extended the BPM to include anisotropy [14].

Two types of incident field illumination were considered:

(i) uniform illumination over the dimensions of the "hard aperture";

and (ii) Gaussian amplitude variation over the aperture, such that the aperture truncated the incident field at the 1/e amplitude points, corresponding to the $1/e^2$ irradiance points, where the irradiance (sometimes called intensity) is given by the square of the modulus of the field amplitude.

The field in the location of the focus was the quantity of

interest. The field was calculated in its complex form, so that the amplitude and the phase of the field could be studied, neglecting the harmonic time-variation. Attention is confined here to the amplitude variation. The field amplitude was calculated over a rectangular grid of 50 axial x 101 transverse points in the vicinity of the focus. An odd number of transverse points were calculated in order to obtain a field symmetric on both sides of the axis. The grid-spacing in the axial direction was either 20 μm or 40 μm, depending on the depth of field. The transverse grid-spacing was 0.6 μm . The area of the focal region calculated, therefore, was 1.0 mm x 60 μm or 2.0 mm x 60 μm.

Representative plots of the field propagating through the focus for both incident conditions (i) and (ii) will now be given. Two aperture widths, 0.6 mm and 3.0 mm will be considered, corresponding to effective f-numbers of f/30.83 and f/6.17 respectively. A great deal of information is available from the computations and the plots are organized as follows:

(a) isometric plots of field amplitude

Figure no.	Incidence conditions
Figure [5.4]	uniform illumination, 2a = 0.6 mm
" [5.5]	Gaussian illumination, 2a = 0.6 mm
" [5.6]	uniform illumination, 2a = 3.0 mm
" [5.7]	Gaussian illumination, 2a = 3.0mm

(b) contour maps of amplitude and encircled energy

Figure no.	Incidence conditions
Figure [5.8]	uniform illumination, 2a = 0.6 mm
" [5.9]	Gaussian illumination, 2a = 0.6 mm
" [5.10]	uniform illumination, 2a = 3.0 mm
" [5.11]	Gaussian illumination, 2a = 3.0mm

The term "encircled energy" is borrowed from bulk optics and denotes the amount of field energy enclosed by circles of

increasing radius from the optical axis. There are no such circles in planar optics, only increasing distances from the axis. The term is a common one, however, and is used here for that reason. The curves shown are normalised to the energy enclosed within a transverse distance of 30 um from the axis at the position where the field is a maximum.

(c) plots of field amplitude along the optical axis of propagation and of the diameter of the central spot, measured along the same axis

Figure no.	Incidence conditions
Figure [5.12]	uniform illumination, 2a = 0.6 mm
" [5.13]	Gaussian illumination, 2a = 0.6 mm
" [5.14]	uniform illumination, 2a = 3.0 mm
" [5.15]	Gaussian illumination, 2a = 3.0mm

The minimum diameter of the central spot along the axis is the waist of the propagating field, and may be considered to locate the focus. Alternatively, the point of maximum field amplitude on the axis may be considered to locate the focus. The two locations need not necessarily coincide, though they do, in general.

(d) plots of transverse enclosed energy at the locations of the axial field amplitude maxima

Figure no.	Incidence conditions
Figure [5.16]	uniform illumination, 2a = 0.6 mm
" [5.17]	Gaussian illumination, 2a = 0.6 mm
" [5.18]	uniform illumination, 2a = 3.0 mm
" [5.19]	Gaussian illumination, 2a = 3.0mm

(e) plots of transverse field amplitude, on linear and logarithmic scales, at locations of axial field amplitude maxima

Figure no.	Incidence conditions
Figure [5.20]	uniform illumination, 2a = 0.6 mm
" [5.21]	Gaussian illumination, 2a = 0.6 mm

"	[5.22]	uniform illumination, 2a = 3.0 mm
"	[5.23]	Gaussian illumination, 2a = 3.0mm

The following discussion will treat the five groups of plots (a) to (e) successively.

5.9.1 Group (a)

The isometric plots of the fields propagating through the focus reveal some general characteristics immediately. Detailed numerical information will be given in later plots. The isometric plots reveal the considerably greater 'smoothness' of the focal field under conditions of truncated-Gaussian incident illumination than under uniform illumination. The peaks and valleys of the uniformly illuminated cases are much more pronounced than those resulting from truncated Gaussian illumination. The truncated-Gaussian incident fields do produce some structure in the focal region, however, as would be expected.

Such behaviour, obtained using the BPM, is consistent with traditional diffraction analyses. The larger field amplitudes of the field outside the central spot for the cases of uniform illumination occur as a result of the Fourier transforming properties of the lens, and indicate the presence of large-amplitude, large-angle spectral components in the input field. Gaussian fields dampen the amplitude of these spectral components, and thereby reduce the height of the sidelobes in the focal region. The damping process can be equivalently obtained by modifying the transmission function of the lens, in which case it is known as 'apodization'.

The different depths of field obtained for different aperture widths are also illustrated in the isometric plots. The narrower apertures give rise to image field amplitudes and beam-diameters that change very slowly with increasing axial distance, whereas the image fields under conditions of larger aperture change rapidly, giving rise to an abrupt focal spot, as is

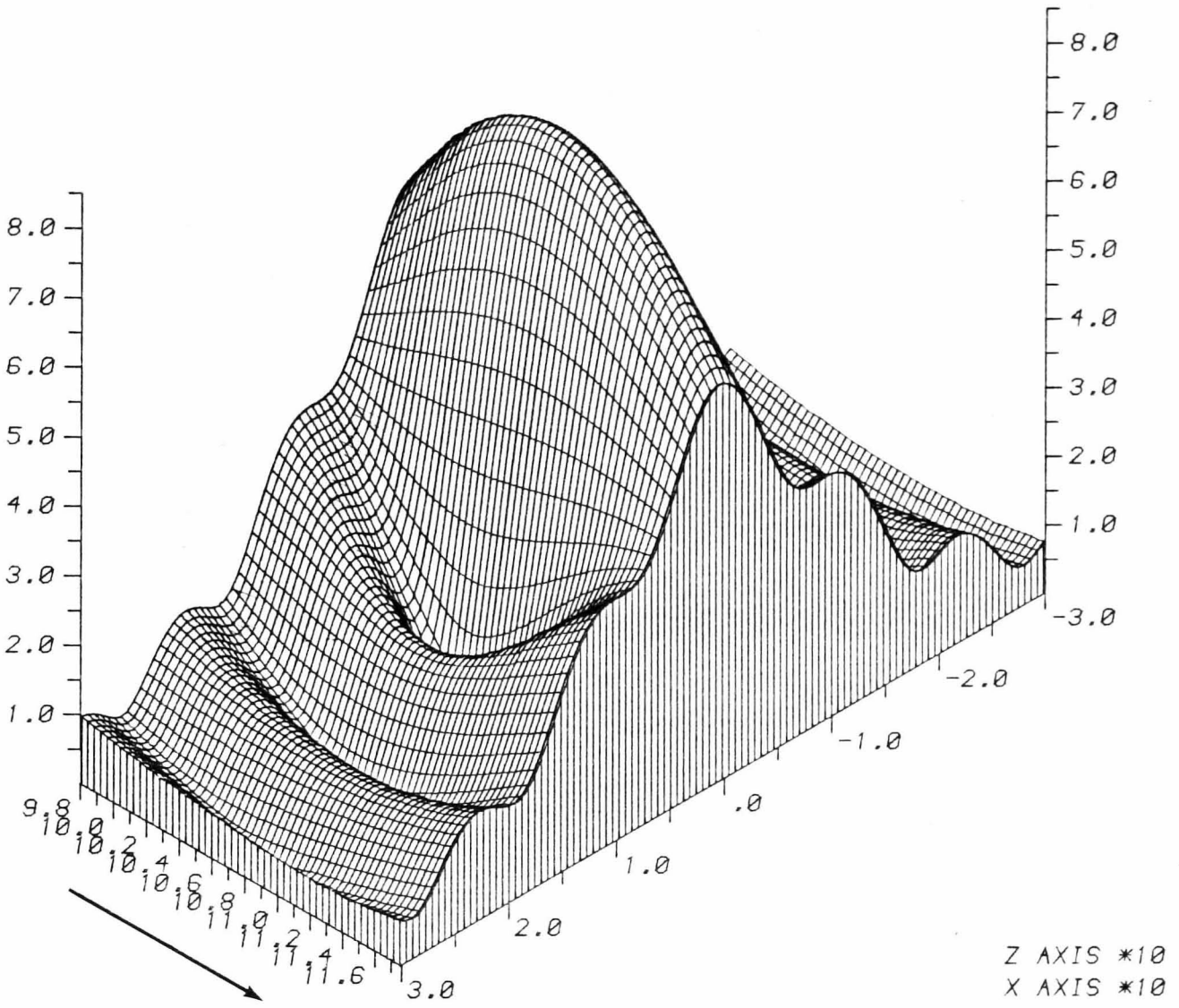


FIGURE [5,4]

Beam propagation through

IOSA Lens; plane wave truncated at 0.6mm

x-axis, μm; y-axis, mm

z-axis, arbitrary units

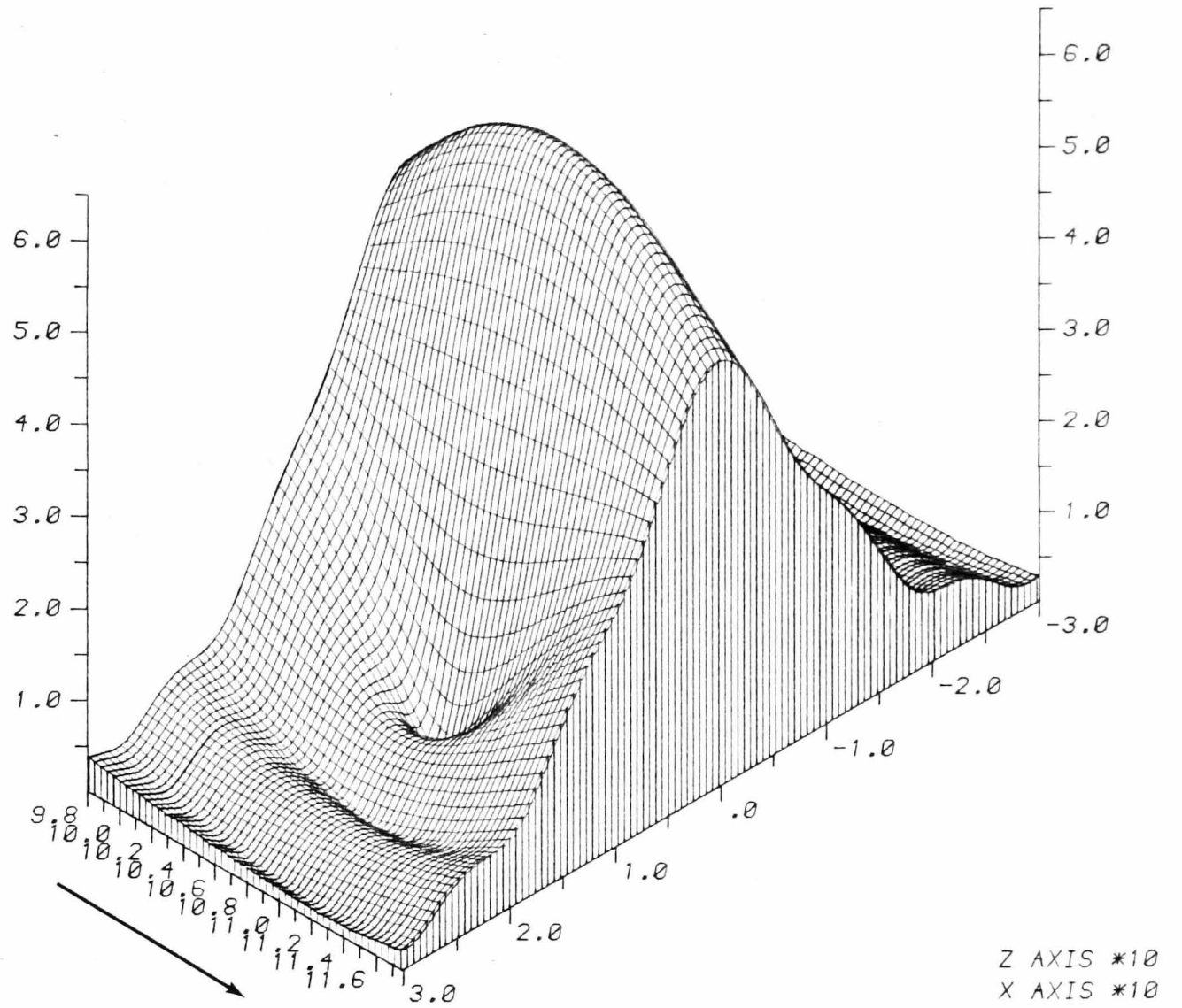


FIGURE [5, 5]

Beam propagation through

IOSA Lens; 0.6mm gaussian truncated at 0.6mm

x-axis, μm; y-axis, mm

z-axis, arbitrary units

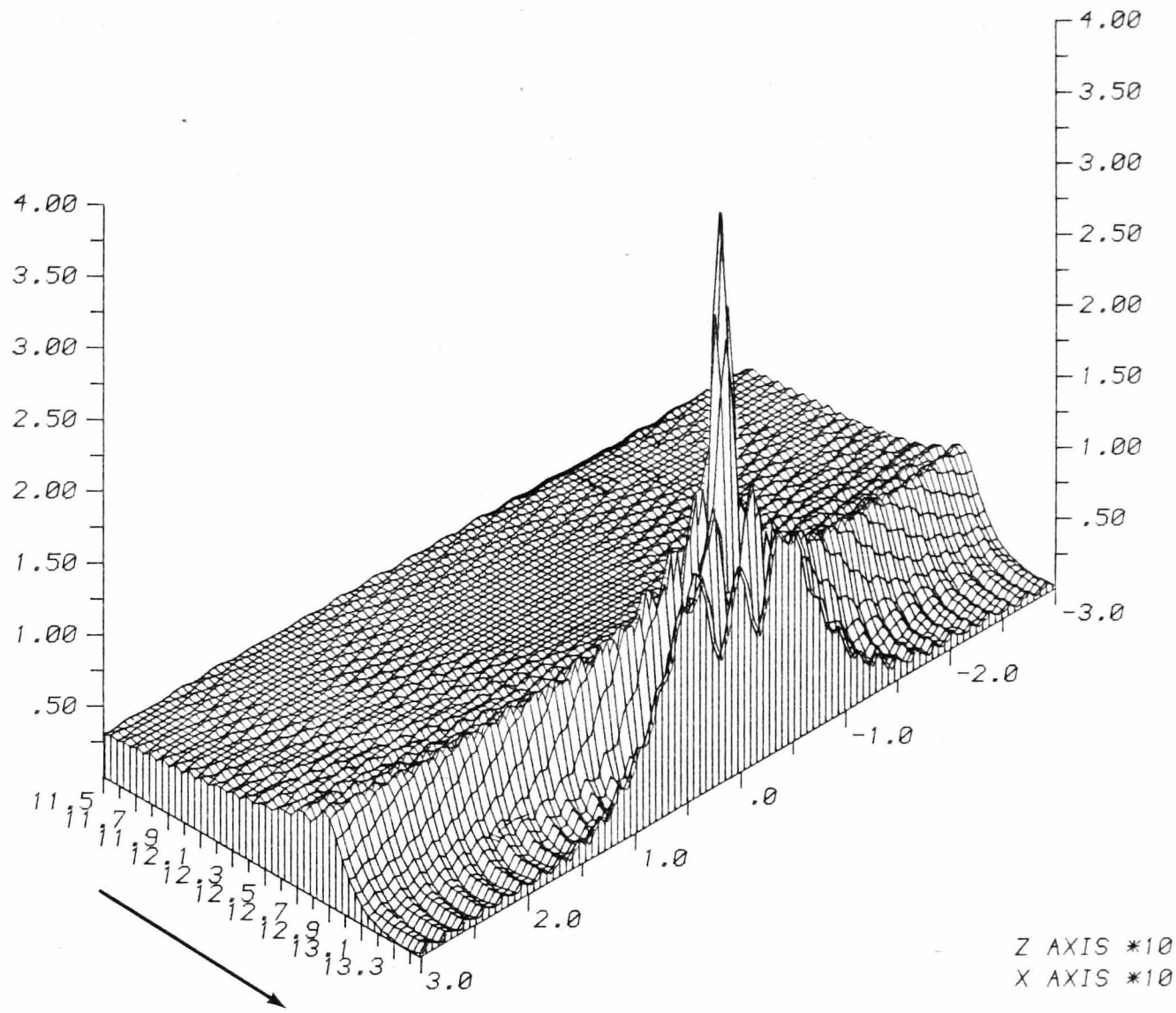


FIGURE [5,6]

Beam propagation through

IOSA Lens; 3.0mm **plane-wave** *truncated at 3.0mm*

x-axis, μm; y-axis, mm

z-axis, arbitrary units

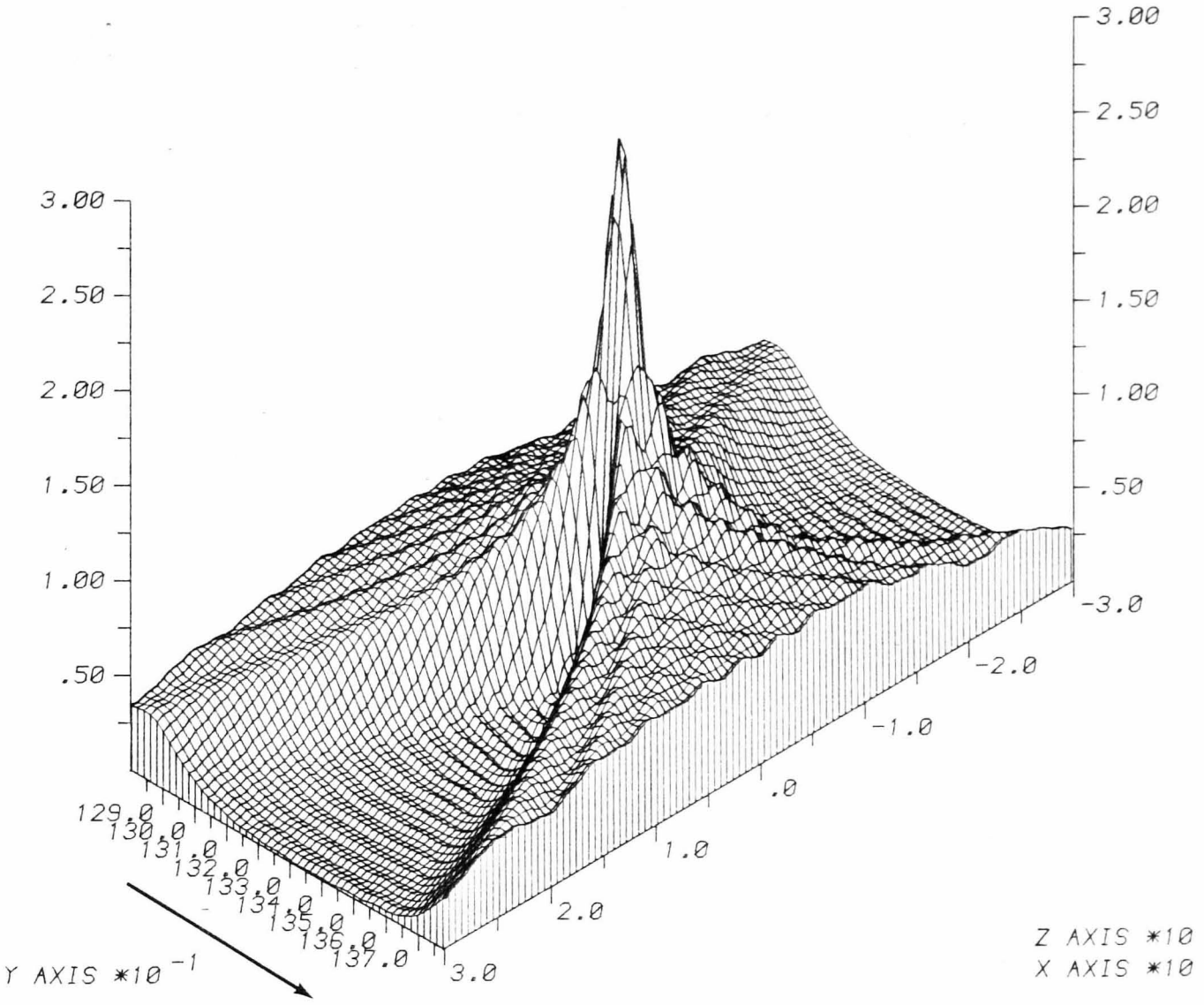


FIGURE [5,7]

Beam propagation through

IOSA Lens; 3.0mm gaussian truncated at 3.0mm

x-axis, μm; y-axis, mm

z-axis, arbitrary units

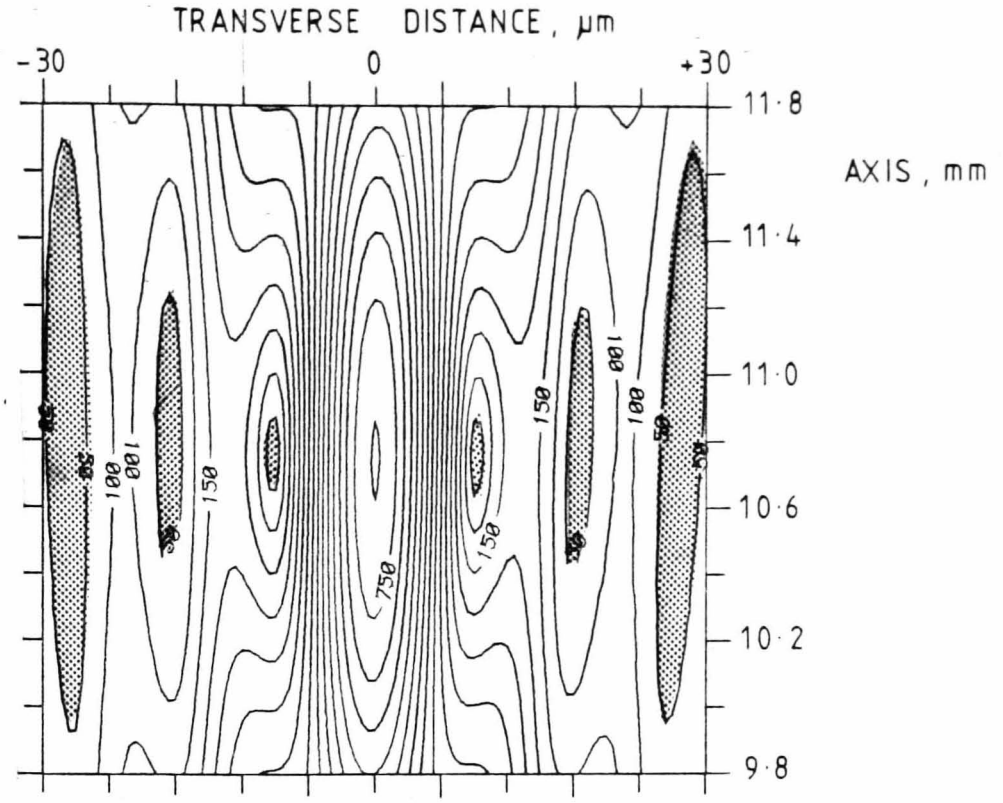
particularly evident in the case of uniform illumination. A large depth of field is obtained, therefore, for narrow apertures whereas a small depth of field is obtained for large apertures, in keeping with the usual cases.

A disconcerting feature of the isometric plots, and, consequently, the plots yet to be discussed, is that the foci do not occur at the distance predicted by geometrical optics (the axial distances shown in the Figures are measured from the exit edge of the lens, and should therefore have 5.0 mm added in order to measure from the centre of the lens). The focal shifts, which are all negative, will be discussed in more detail later.

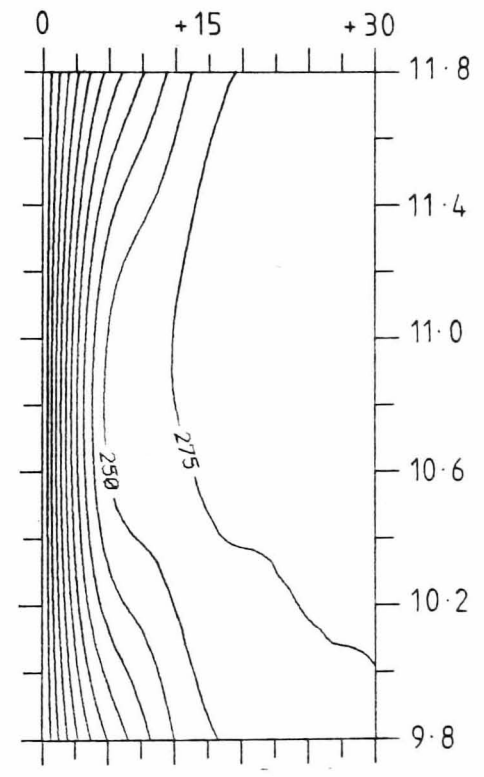
5.9.2 Group (b)

The contour maps are more informative than the isometric plots. Maps of both equi-amplitude contours and equi-energy contours are given for each case of incident illumination. The maps are similar to those presented by Born and Wolf [12]. As pointed out by them, the equi-energy contours may be considered analogous to the rays of geometrical optics.

For the case of narrow, uniform illumination, a distinct central field-amplitude structure is observed at the focus, surrounded by sidelobes. For the circularly symmetric bulk optical case, Born and Wolf described the central structure of the focal field, under conditions of uniform illumination, as 'tubular', since a tube of light occurs which is circumscribed in all directions by the first minimum in the diffraction pattern. The amplitude pattern is symmetrical about the optical axis, and looks symmetrical about the transverse line through the focus. The encircled energy contours are not symmetrical about this transverse line, however, especially the contours furthest from the axis. Such asymmetry is not predicted by classical diffraction analyses [12], although recent work has indicated that asymmetries may indeed arise. The encircled energy contours do indeed have the appearance of geometrical rays, except for the highest-energy contours furthest from the axis.



(a)

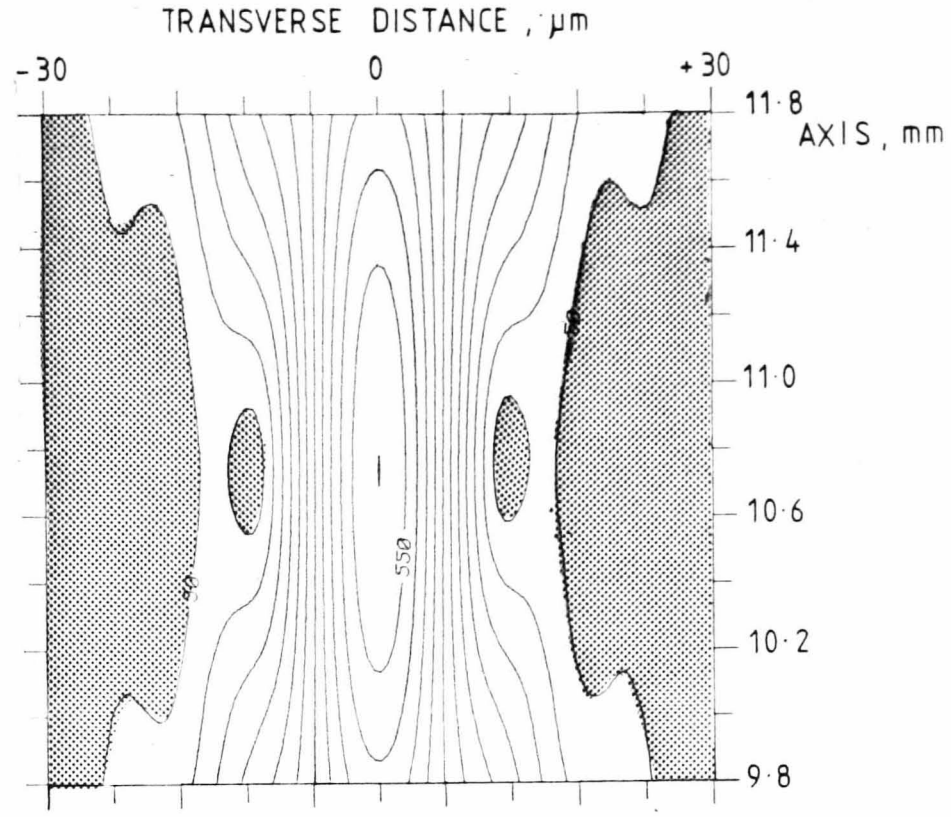


(b)

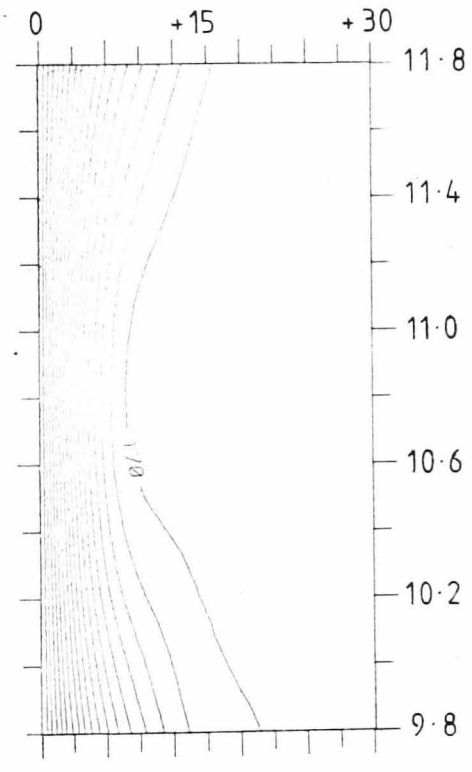
FIGURE [5, 8] *Contour maps*

IOSA Lens; plane wave truncated at 0.6mm

(a) Field
(b) encircled energy



(a)

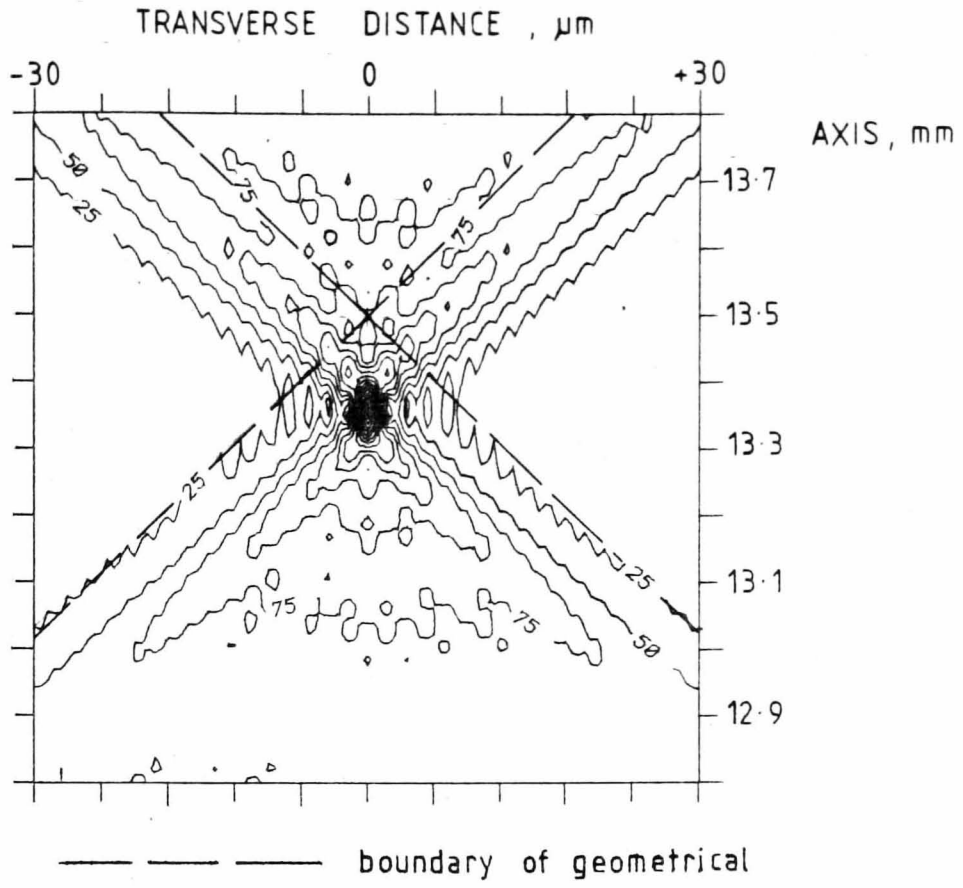


(b)

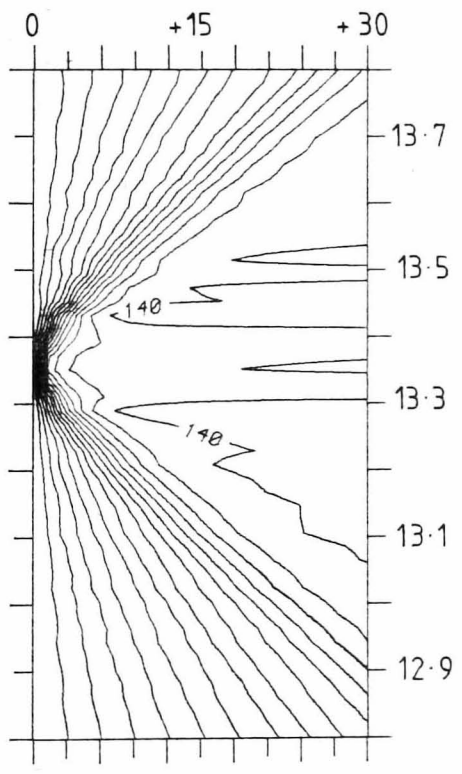
FIGURE [5, 9] Contour maps

IOSA Lens; 0.6mm gaussian truncated at 0.5

(a) Field
(b) encircled energy



(a)

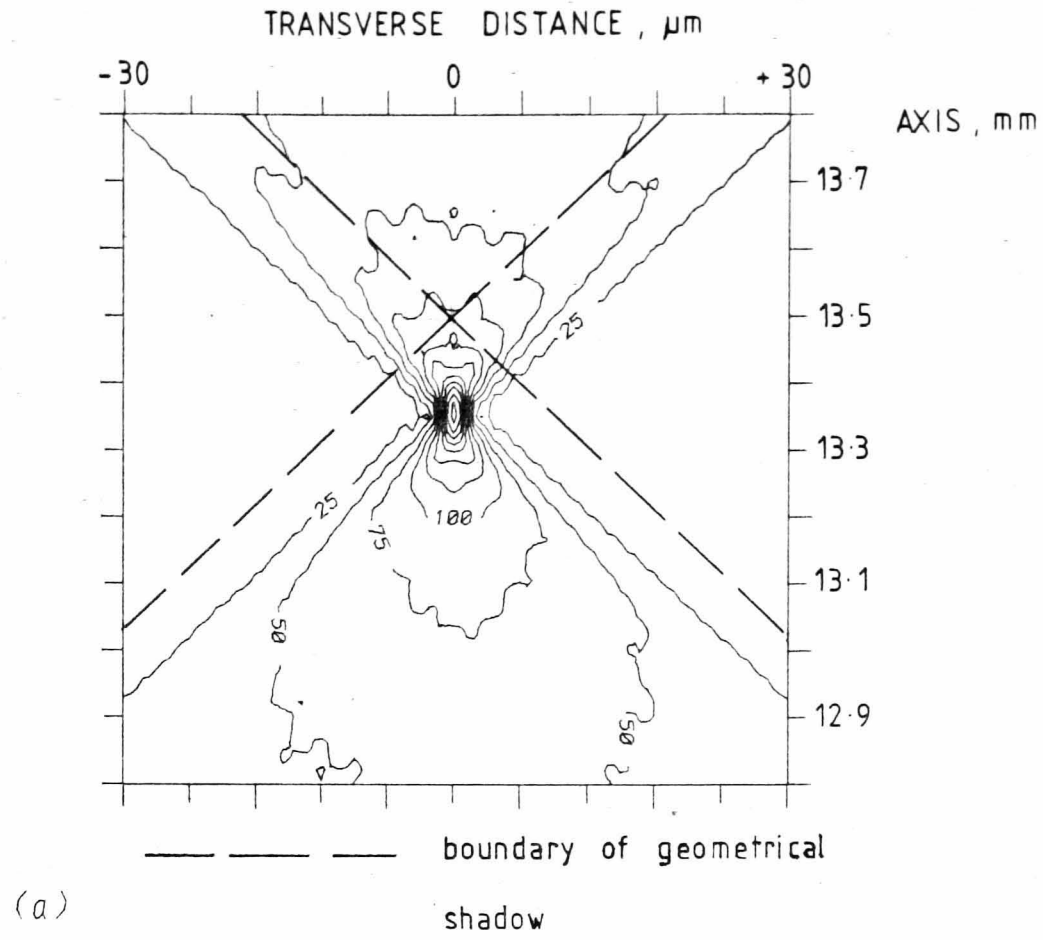


(b)

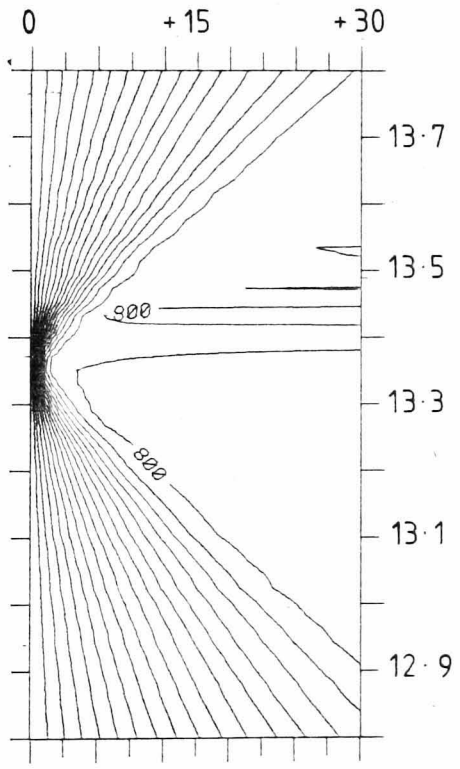
FIGURE [5,10] *Contour maps*

IOSA Lens; 3.0mm 'plane' wave

(a) Field
(b) encircled energy



(a)



(b)

FIGURE [5,11] *Contour maps*

IOSA Lens; 3.0mm gaussian truncated at 3.0mm

(a) Field
(b) encircled energy

As expected, the case of narrow, truncated-Gaussian illumination is somewhat similar to the case of narrow, uniform illumination. However, the tubular structure centred on the focus is not so distinctive indicating that the sidelobes are suppressed. The focal field does not rise as sharply as in the uniformly-illuminated case, as indicated by the density of contour lines.

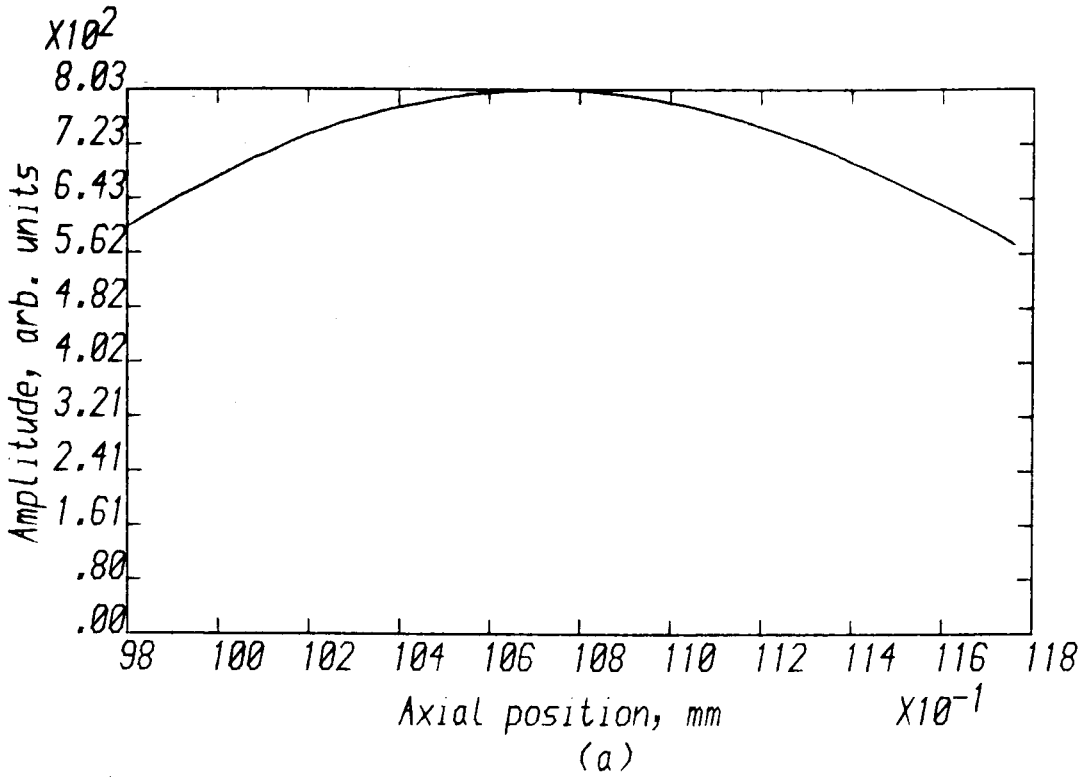
The contours enclosing the regions of lowest amplitude have been shaded for both cases of narrow incident fields, which helps to illustrate the wider extension of the focal field for the case of uniform illumination.

The case of broad, uniform illumination gives rise to a very intense focal spot, again having a tubular structure in the close vicinity of the focus. Far from the focus, the amplitude contours display a four-pointed star shape. Such structure indicates the considerable contribution made to the diffraction pattern by the field at the boundaries of the diffracting aperture. The boundary of the geometrical shadow is shown, and it is again apparent that the focal distance calculated using the BPM is not the one predicted by geometrical optics. The equi-energy contours display considerable distortion near the focus.

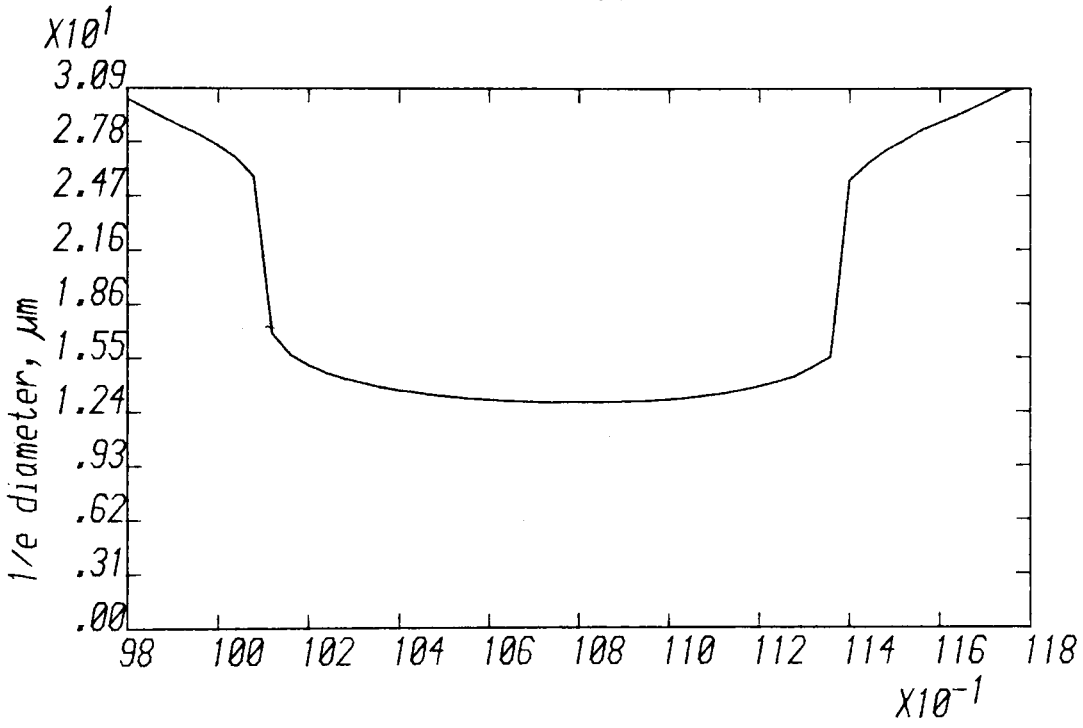
In contrast, the case of broad, truncated-Gaussian illumination gives rise to equi-amplitude contours that are tubular near the focus and figure-of-eight shaped far from the focus. The star-structure is absent. The equi-energy contours are also much smoother. Once more, the focus calculated by the BPM is not the one predicted by geometrical optics.

5.9.3 Group (c)

The axial amplitude and 1/e-diameter plots yield the actual values calculated for the foci for different illumination conditions.



(a)



(b)

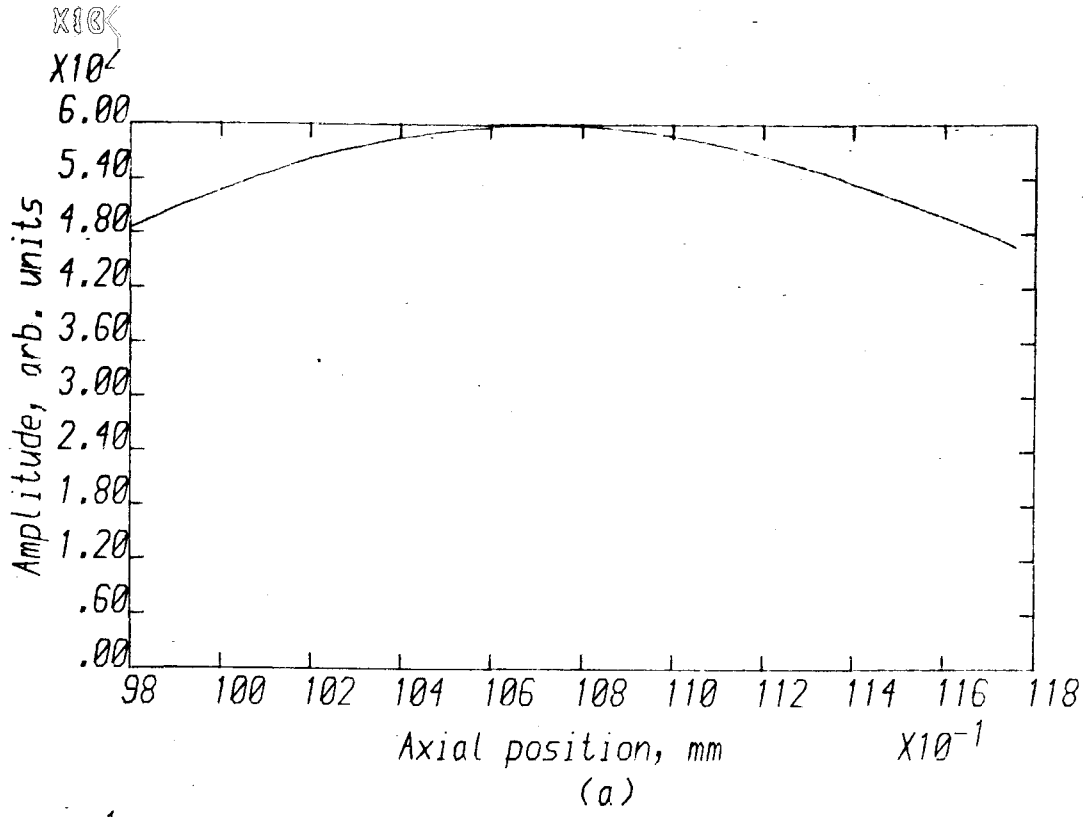
FIGURE [5,12]

Axial profiles: 0·6mm plane wave

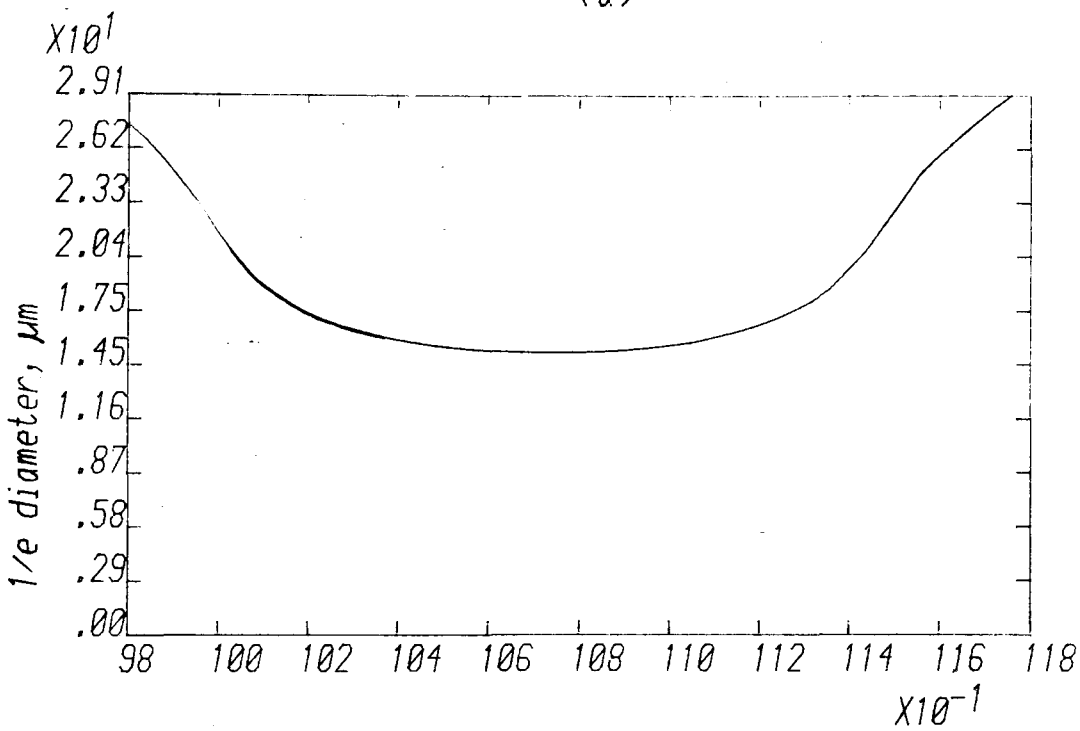
(a) axial Field

maximum Field at z= 10.72 mm
waist diameter = 12.92 µm
(b) 1/e beam diameter

waist at z= 10.80 mm
waist diameter = 12.90 µm



(a)



(b)

FIGURE [5, 13]

Axial profiles: **0·6mm gaussian**

(a) axial field

maximum field at z= 10.72 mm
waist diameter = 15.20 µm

(b) 1/e beam diameter

waist at z= 10.72 mm
waist diameter = 15.20 µm

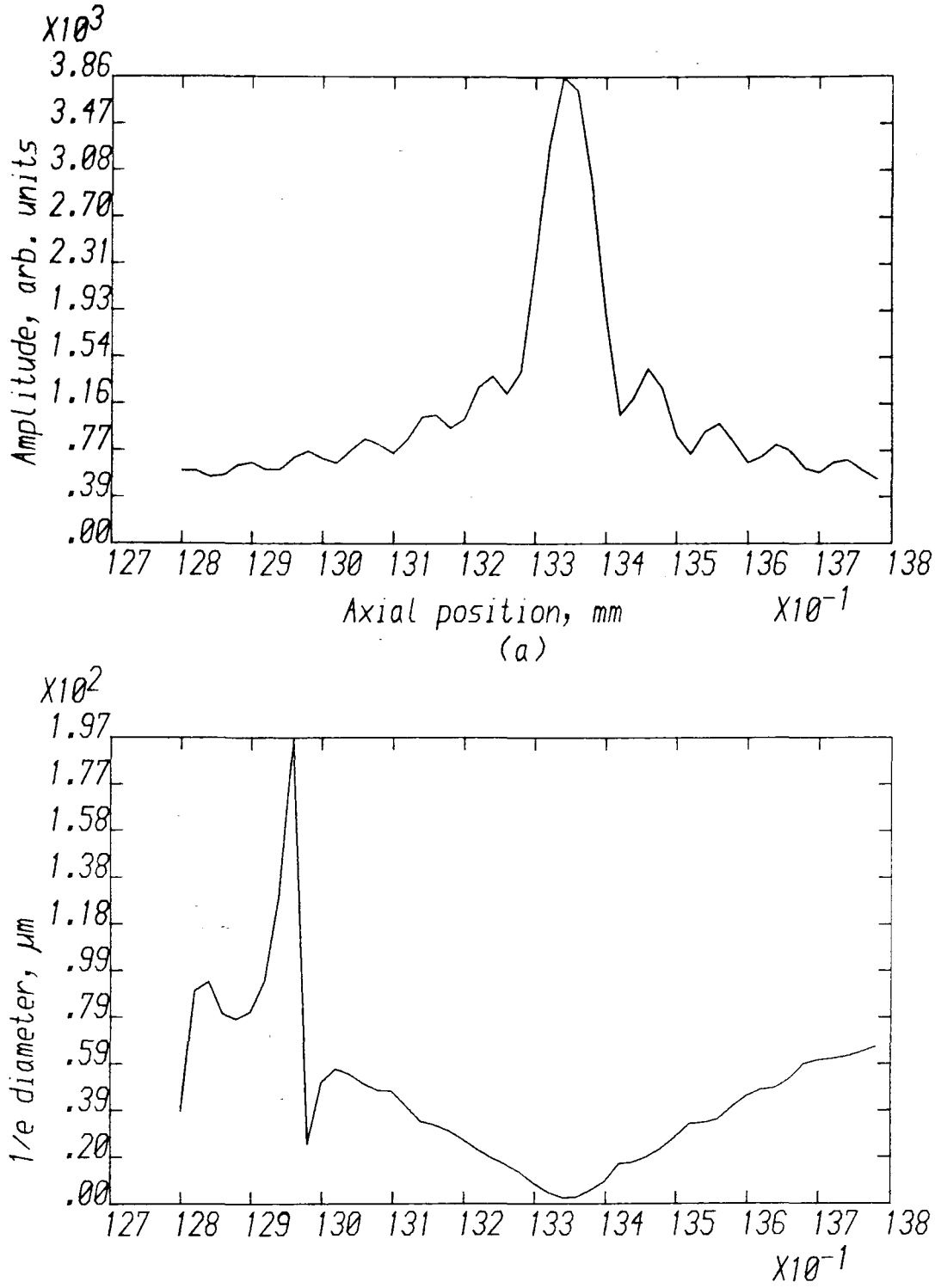


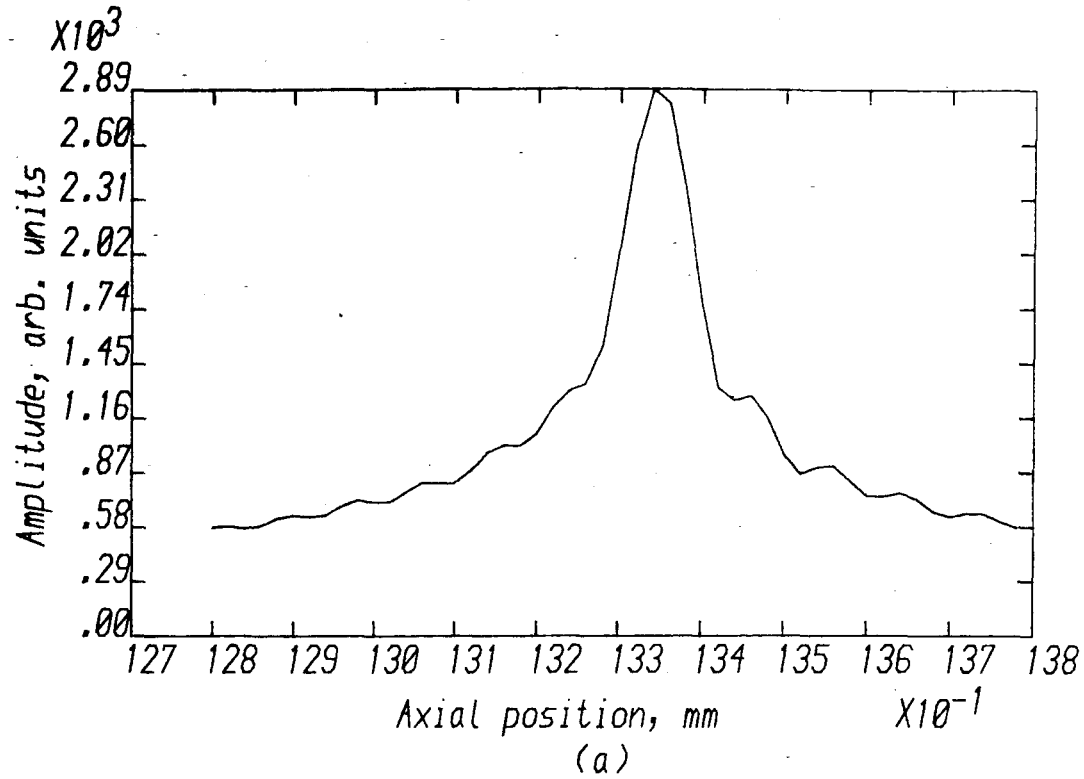
FIGURE [5,14]

Axial profiles: 3·0 mm plane wave

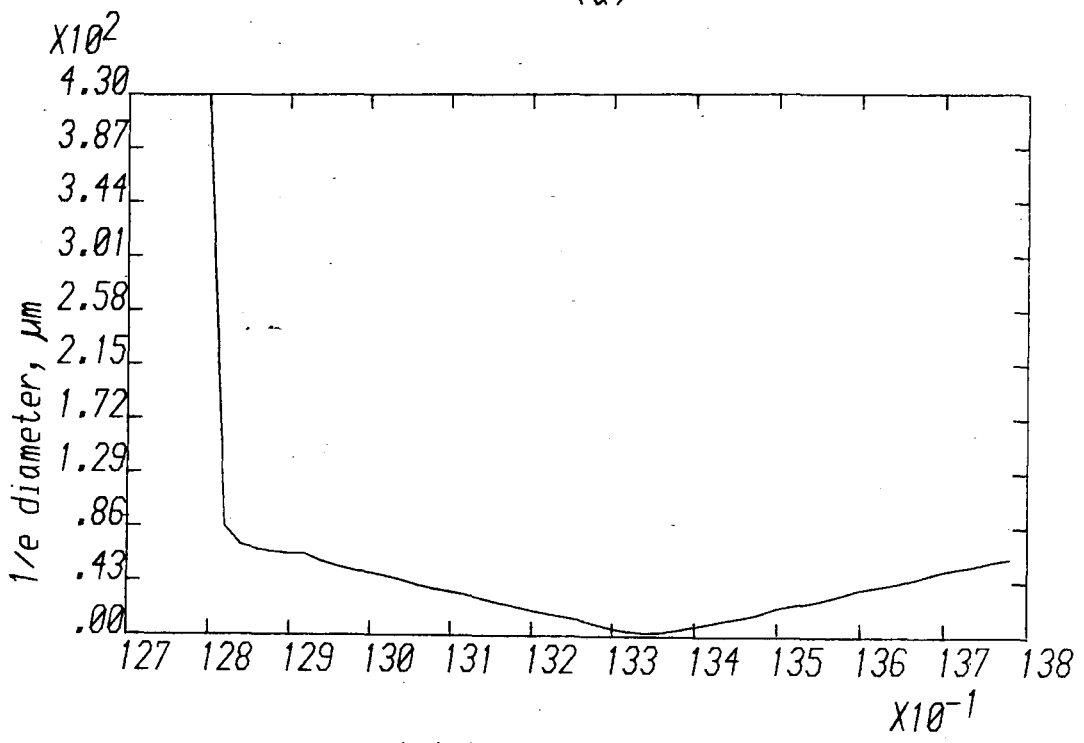
(a) axial Field

maximum Field at z= 13.34 mm
waist diameter = 2.55 µm
(b) 1/e beam diameter

waist at z= 13.34 mm
waist diameter = 2.55 µm



(a)



(b)

FIGURE [5, 15]

Axial profiles: **3·0mm gaussian**

(a) axial field

maximum field at z= 13.34 mm
waist diameter = 2.98 µm
(b) 1/e beam diameter

waist at z= 13.34 mm
waist diameter = 2.98 µm

For narrow, uniform illumination the axial amplitude maximum occurs at 10.72 mm from the exit edge of the lens, ie 15.72 mm from the centre of the lens. The beam-diameter at this point is 15.20 μm. The waist of the beam does not coincide with the axial amplitude peak and occurs at 15.8 mm from the lens centre. However, the waist diameter is 15.20 μm, as for the peak amplitude case, showing that the non-coincidence of amplitude maximum and waist is due to the large depth of field. The variation of the focal field width will be considered in detail later.

The 1/e-diameter curve displays two step-like changes. The steps indicate that the central lobe has decreased in amplitude whilst the sidelobes have increased. Essentially, the propagating field is losing the one-dimensional Airy function shape [12] and re-acquiring a flat-topped shape.

The behaviour of the axial amplitude curve in the case of narrow, truncated-Gaussian illumination is similar to that for uniform illumination. The maximum values occur at the same distance, 15.72 mm. The absolute maximum value of the amplitude, 600 arbitrary units, is less than that of the uniformly illuminated case, 803 arbitrary units, as expected, since a uniform field carries more energy than does a truncated Gaussian of the same width and peak amplitude. The beam waist coincides with the maximum of the field and is 15.20 μm. The beam diameter curve does not, however, display step-like changes, indicating smooth propagation.

Broad, uniform illumination gives rise to an axial diffraction pattern similar to the transverse patterns. A number of local maxima and minima occur, together with the focal peak which is the absolute maximum and which occurs at a distance of 18.34 mm from the centre of the lens. The focal peak is quite narrow in the axial direction, having a half-power width of just over 100 μm. The beam-diameter curve fluctuates until it approaches the focal region where a clearly-defined central lobe is present. The curve then decreases to a clearly defined minimum

at 18.34 mm from the centre of the lens. The waist is 2.55 μm wide. At the design focal-length of 18.5 mm, the beam width is approximately 30 μm and the amplitude is down to only 22% of the peak value.

Broad, truncated-Gaussian illumination also gives rise to a sharp focal peak. However, other local minima and maxima are not so much in evidence. The beam-diameter curve is smooth through most of its length. The focal length is once more 18.34 mm from the centre of the lens, though the waist diameter at 2.98 μm is slightly larger than that for the uniformly-illuminated case.

5.9.4 Group (d)

The encircled energy curves provide for further useful quantitative comparisons between the cases of truncated-Gaussian and uniform illumination.

At the narrow aperture, the uniformly illuminated case and the truncated-Gaussian case both give rise to similar encircled energy curves over the first 90% or so of energy, indicating the similarity between the central lobe of a sinc curve and a Gaussian of approximately the same width. Both curves show that nearly 90% of the energy in the focal field is contained within 6 μm either side of the axis. However, the truncated-Gaussian case shows more favourable behavior in concentrating the remainder of the energy,with 98% of the energy concentrated within 12 μm of the axis, in comparison with the figure of 21 um for the uniformly illuminated case.

At broad apertures, truncated-Gaussian illumination is even more favourable, with 98% of the energy concentrated within 3 μm of the optical axis in comparison with 12 μm for the uniformly illuminated case.

5.9.5 Group (e)

Finally, the transverse amplitude profiles are shown for

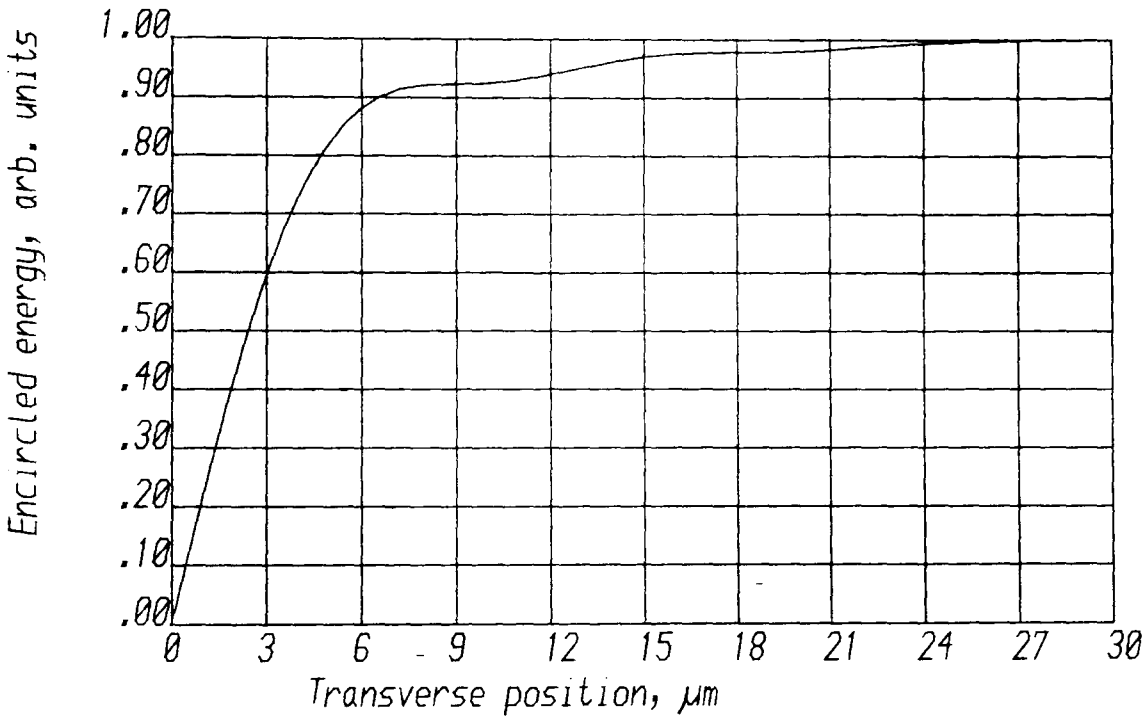


FIGURE [5,16] 0·6 mm plane wave

Encircled energy of transverse field

at axial distance, z= 10.72 mm

values are here normalised to energy at

plane of maximum amplitude

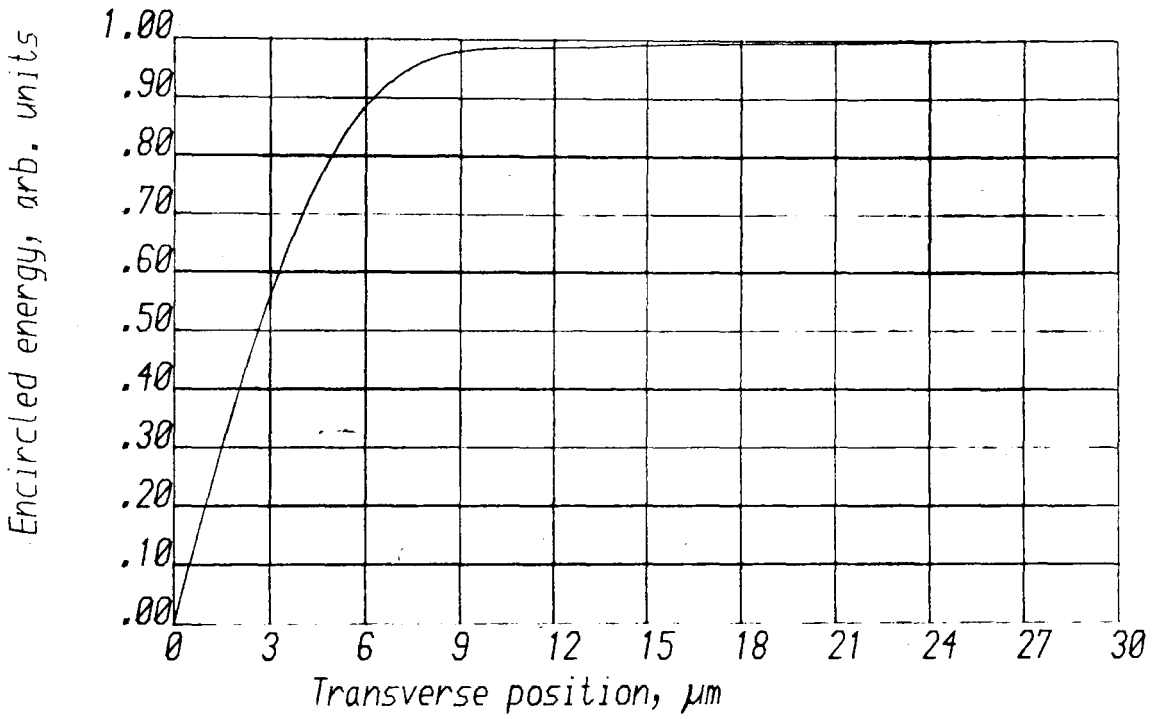


FIGURE [5,17] 0·6 mm gaussian

Encircled energy of transverse field

at axial distance, z= 10.72 mm

values are here normalised to energy at

plane of maximum amplitude

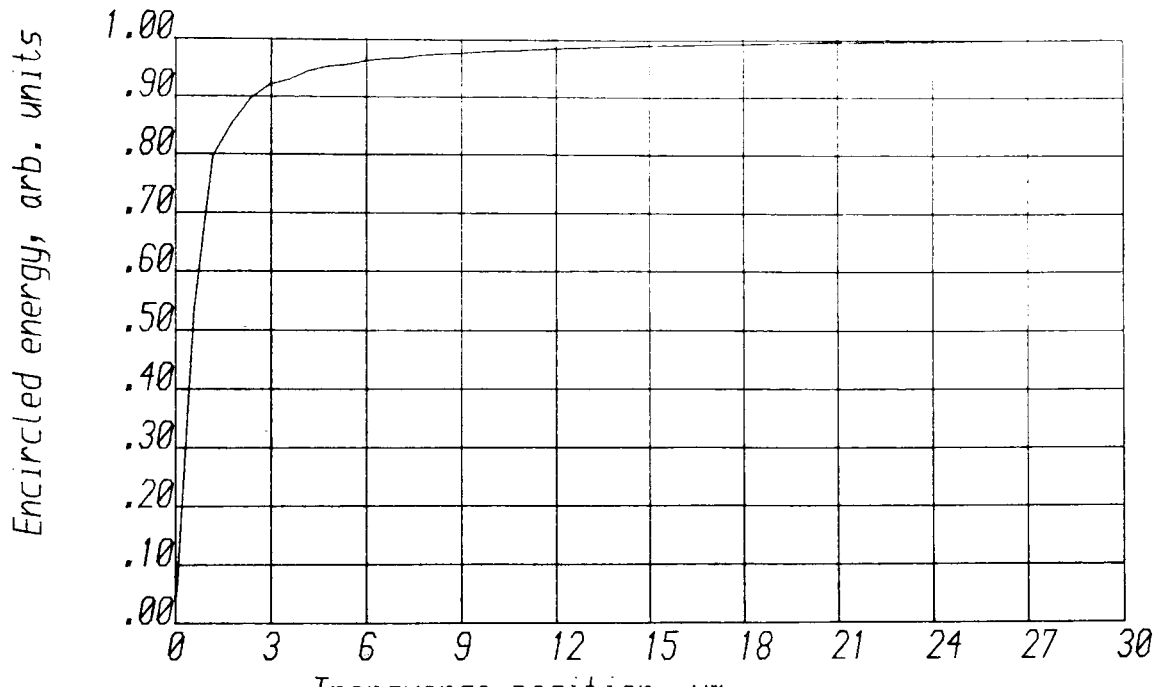


FIGURE [5,18] 3·0 mm plane wave

Encircled energy of transverse field

at axial distance, z= 13.34 mm

values are here normalised to energy at

plane of maximum amplitude

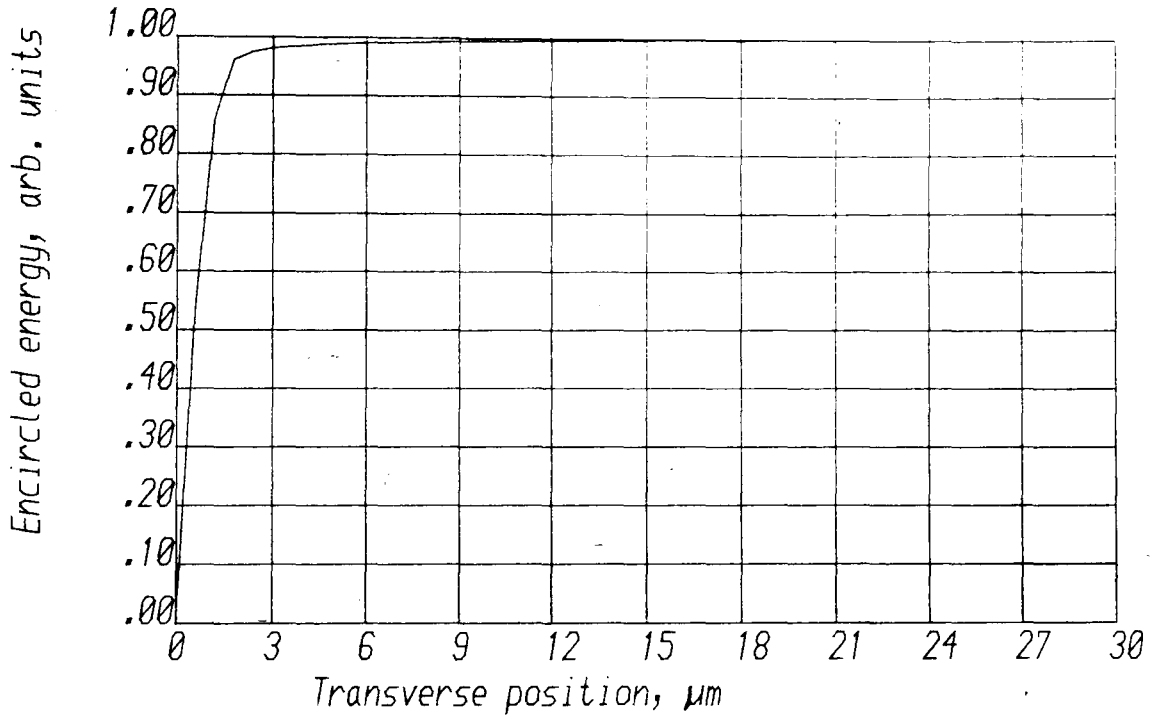


FIGURE [5,19] 3·0mm gaussian

Encircled energy of transverse field

at axial distance, z= 13.34 mm

values are here normalised to energy at

plane of maximum amplitude

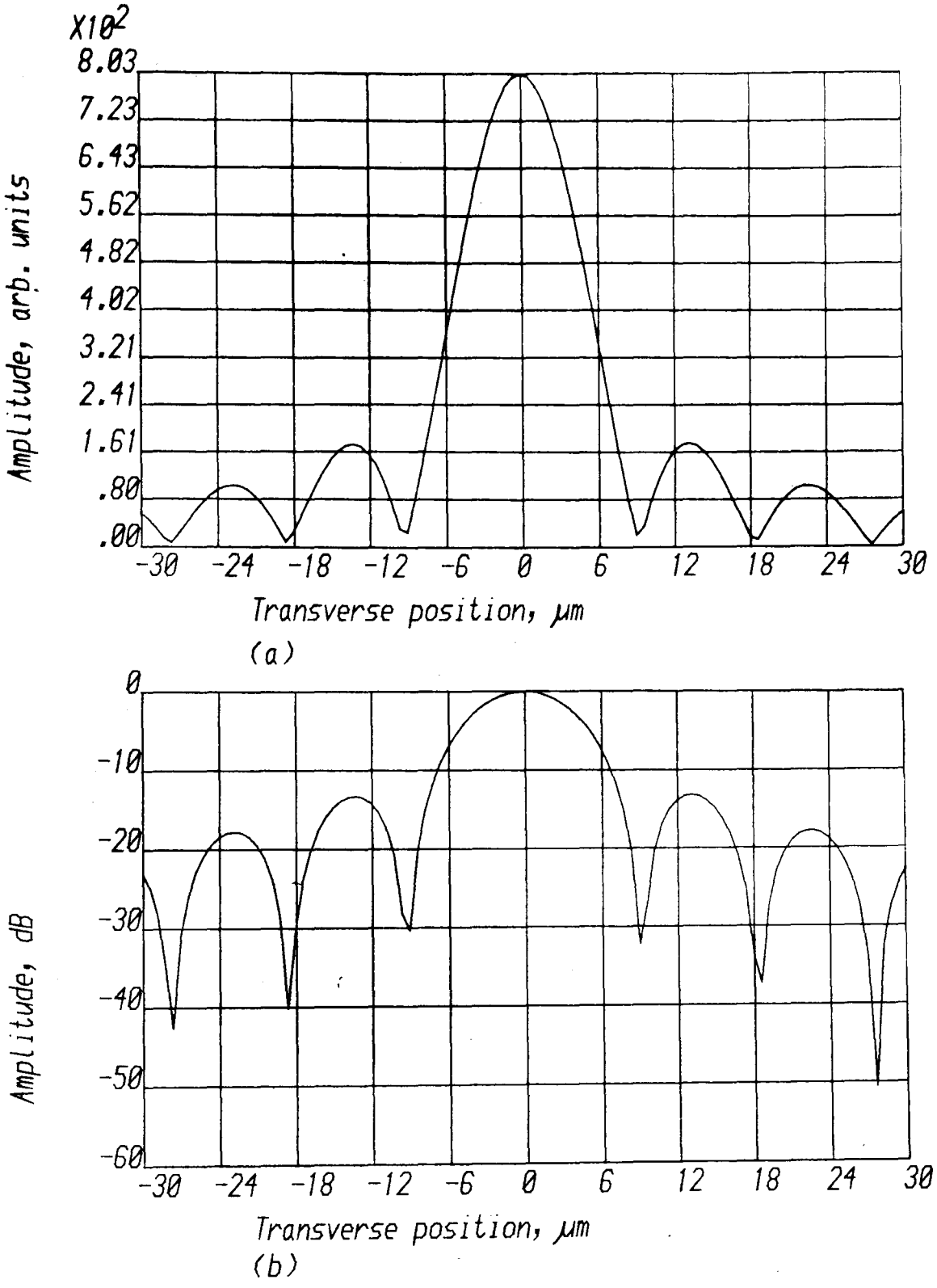


FIGURE [5,20] *Transverse Field*

at axial distance, z= 10.72 mm

(a) linear

(b) logarithmic

0·6mm plane wave

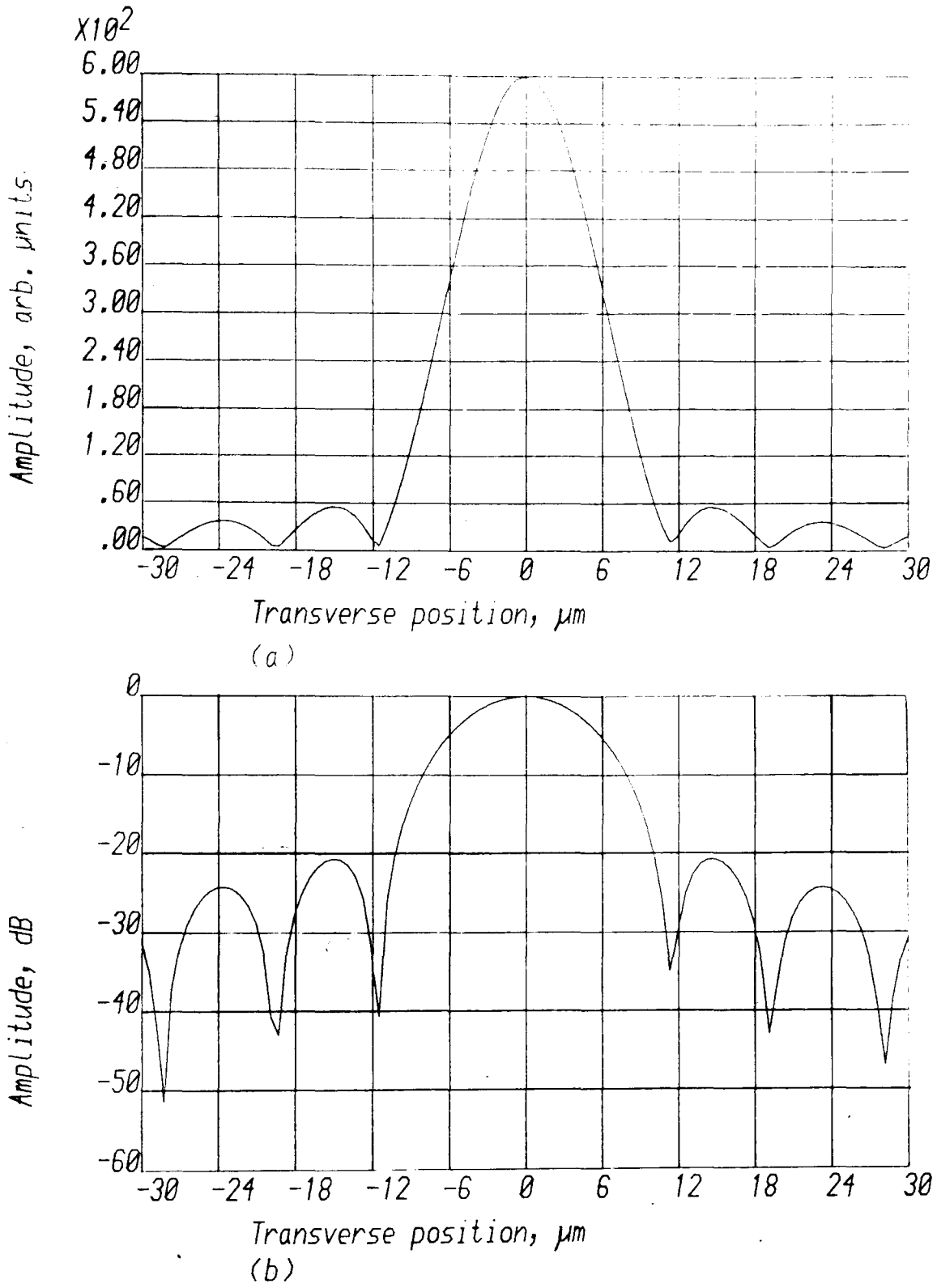


FIGURE [5, 21] *Transverse Field*

at axial distance, z= 10.72 mm

(a) linear

(b) logarithmic

0·6 mm gaussian

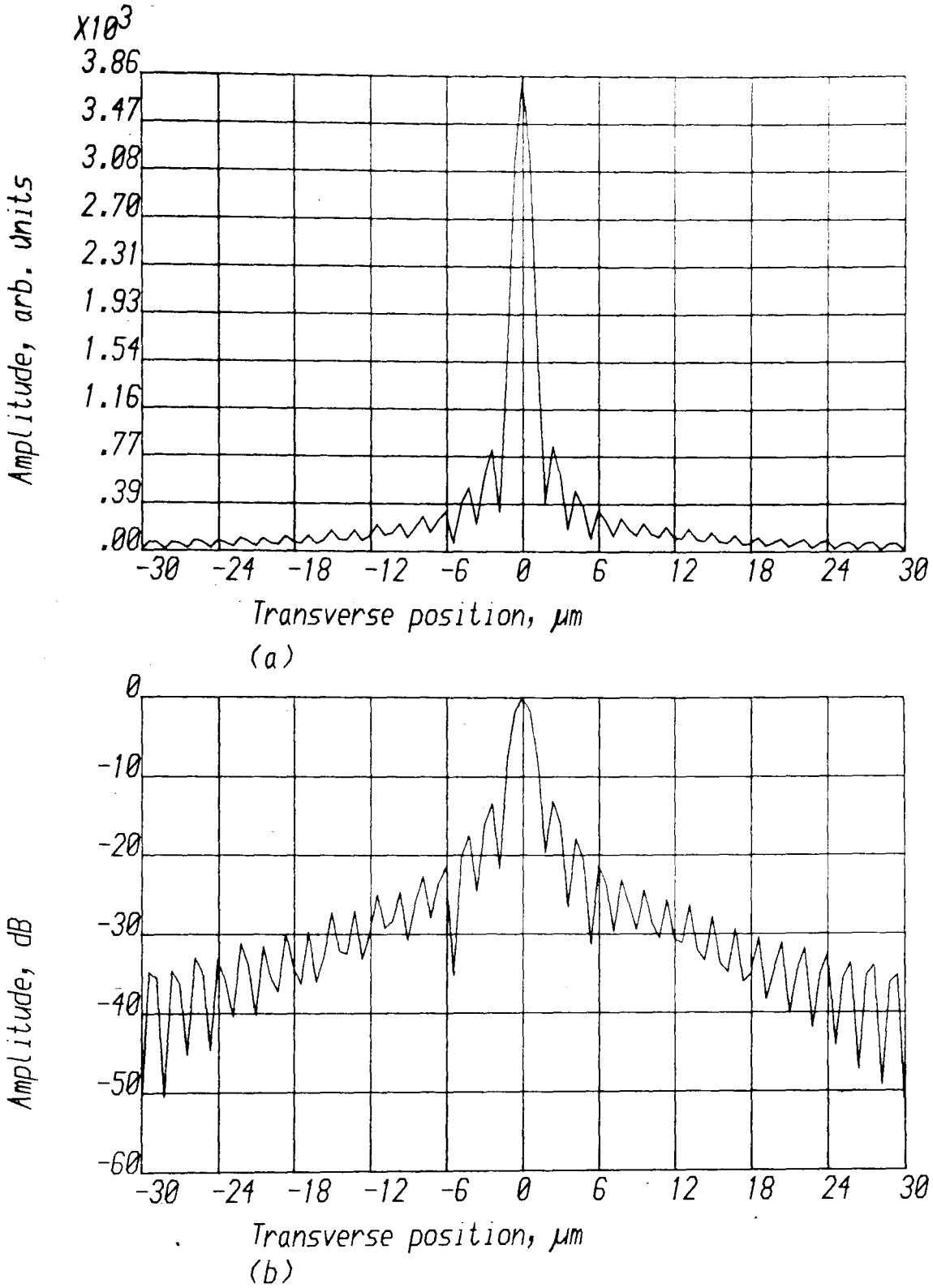


FIGURE [5, 22] *Transverse Field*

at axial distance, z= 13.34 mm

(a) linear

(b) logarithmic

3·0 mm plane wave

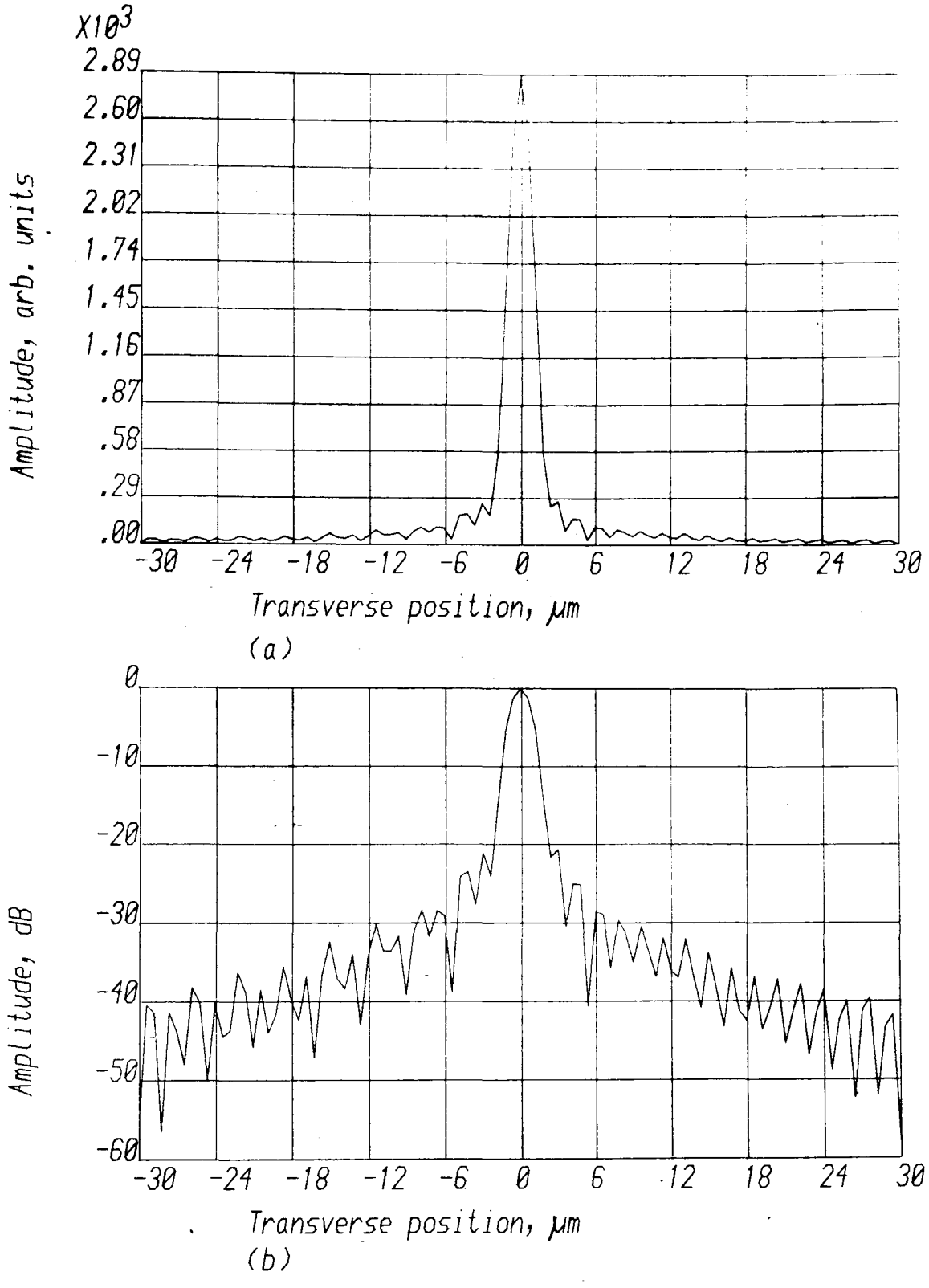


FIGURE [5,23] *Transverse Field*

at axial distance, z= 13.34 mm

(a) linear

(b) logarithmic

3·0mm gaussian

each case of illumination. Well-defined sinc functions are obtained for the case of uniform incident illumination. Similar functions are obtained for the case of truncated-Gaussian illumination, but with the sidelobes significantly reduced, the first sidelobe being more than 20 dB below the maximum value for both narrow and broad illumination.

5.10 Discussion

The focal-shift observed in the above Figures requires some explanation. Figure [5.24] shows the normalised focal-shift observed as a function of the width of the truncating aperture, for both uniform illumination of the aperture and Gaussian illumination. The curves are exactly the same for both types of illumination, indicating that the shift depends on the aperture width only, and not on the amplitude-profile of the beam illuminating the aperture. However, it might be expected that Gaussian beams with very weak truncation would display different characteristics. Such conditions were not investigated. The results were found to be stable to variations in the sampling interval within the allowed band of Figure [8.2], to variations in the number of discrete points at which the lens profile was specified, to variations in the propagation step and to variations in the number of samples in the guard-bands . These checks helped to establish that the results were not artefacts of the BPM itself.

Focal-shifts have been widely reported in the literature [15-19]. The key features of the shift are :

(i) that a focal-shift always results in the true focus being located closer to the lens than the geometrical focus;

(ii) that the shift is largest for small angular apertures, ie for large effective f-numbers.

Both of these characteristics were present in the results

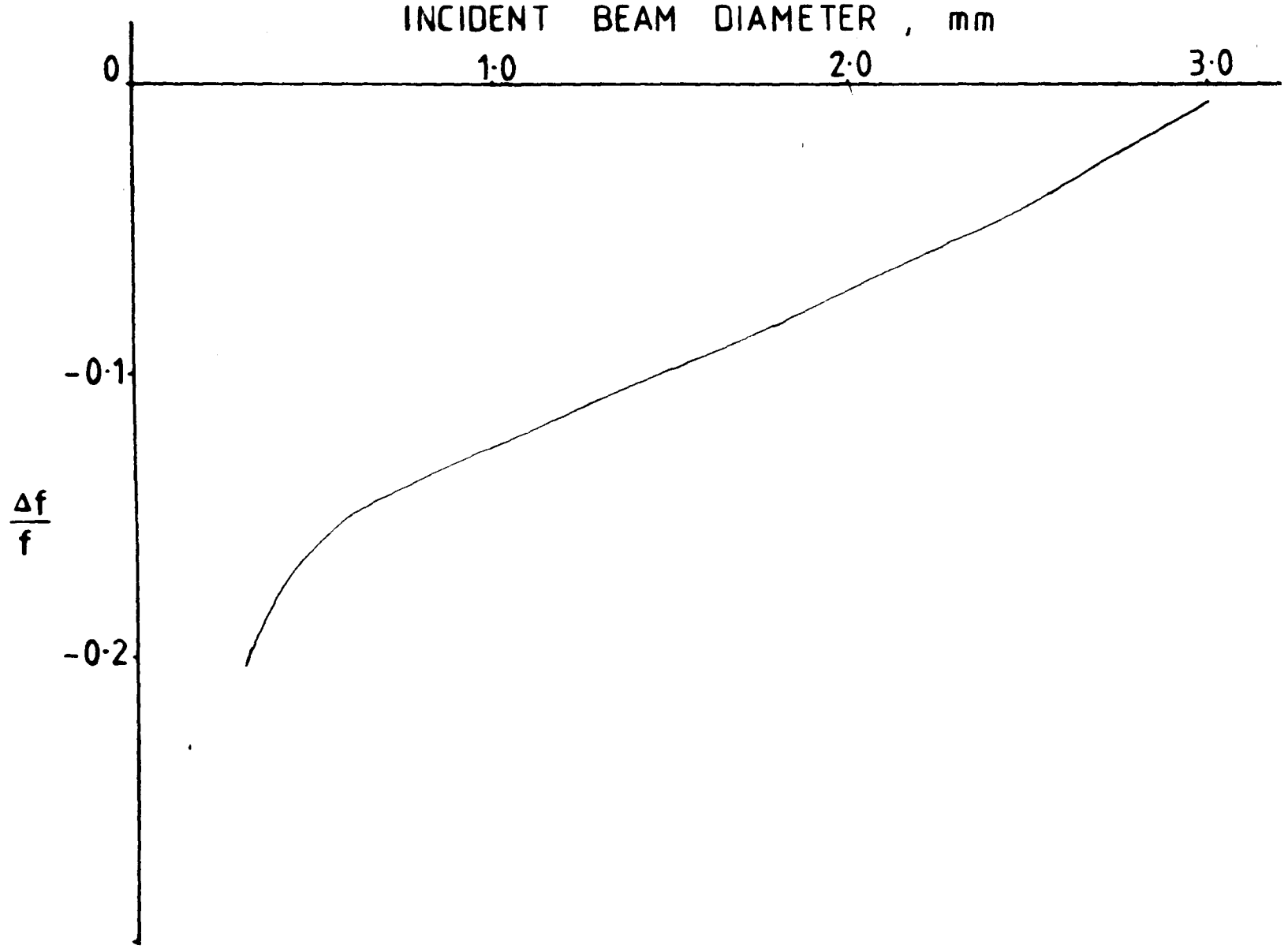


FIGURE [5, 24] Focal shift observed for both types of incident field

computed above. The theoretical calculations presented in references [15-19] were carried out for thin-lenses having circular symmetry, immersed in air. It might be expected that the results would be different for thick planar lenses situated on a high-index substrate, such as the one investigated in the present chapter. The presence of the edge-rounding region in the geodesic lens could also influence the focal-field patterns, as has been observed by van der Donk [14]. However, it is worthwhile to compare quantitative results with those reported in the literature. An important number is the so-called Fresnel number of the aperture, given by:

$$N_a = \frac{a^2 n}{\lambda f} \qquad (5.45)$$

when viewed from the geometrical focus. The quantities on the right-hand side have already been defined in this chapter. The focal-shift obtained for an aperture 2a of 0.3 mm may be investigated. The lens is then operating at the fairly large f-number of f/61.7 . Using f = 18.5 mm, λ = 0.633 μm and n = 2.2065, the Fresnel number of the aperture is then 4.24 . If the truncated Gaussian case is considered, a Fresnel number N_w may also be defined for which a in equation (5.45) is replaced by w, the Gaussian waist radius. Since w = a in the case considered, $N_w = N_a = 4.24$. Li and Wolf [15] have published universal curves from which the focal-shift may be estimated from a knowledge of N_w and N_a. For $N_w = 4.24$, the focal-shift should be between 6% and 8%. The value obtained in the present case was 20% . If it is allowed that the Fresnel number of the aperture takes into account the further 5.0 mm that the wave-field would have to travel from the aperture to the centre of the thick geodesic lens, ie if f = 23.5 mm is allowed, then the Fresnel number becomes 3.34 . In such a case, the focal-shift predicted by Li and Wolf would be greater than 10% . Furthermore, the percentage focal-shift obtained using the BPM would then be recalculated, using f = 23.5 mm, as 15.7% . Saga et al [17] have also published curves of focal-shift as a function of the truncation ratio a/w and a parameter P defined by:

$$P = \frac{kw^2}{f} \qquad (5.46)$$

where k is the wave-number in the material. If f = 23.5 mm is used in this case, then P = 20.99 . The focal-shift estimated from the curves of Saga et al would then be approximately 15% . The results presented for narrower apertures in Figure [5.24] are, therefore, plausible. At wider apertures the focal-shift is surprisingly large, being 160 μm for an input beam-width of 3.0mm. If such a shift indeed exists, the implications for lens-design are quite obviously disturbing. More work is required to determine the accuracy of these results.

Parker-Givens [18] has given a physical interpretation of the reasons for the presence of a focal-shift in the diffracted field. The distance s from the point of observation on the axis to different points on the aperture is usually treated as a constant value, namely the geometrical focus f , in conventional theories of diffraction. If the distance s is treated as a variable, however, the results of the diffraction calculation are considerably altered, especially at small angular apertures. In bulk optics, with spherical lenses, the result is that an inverse square law acts in competition with the constructive interference caused by the progression of the wavefront to the geometrical focus, generating a focal-shift.

The BPM generates near-diffraction limited field profiles, as shown in Figure [5.25]. The theoretical 1/e beam-diameters as a function of beam-width for both Gaussian and uniform illumination are plotted, together with the results obtained using the BPM. An untruncated Gaussian beam having a 1/e-beam-diameter equal to a uniformly illuminated aperture will generate a smaller 1/e diffraction spot than the uniformly illuminated aperture, since the full aperture of the untruncated beam extends to infinity, theoretically. However, a Gaussian beam truncated at 1/e points equal to the width of a uniformly illuminated aperture

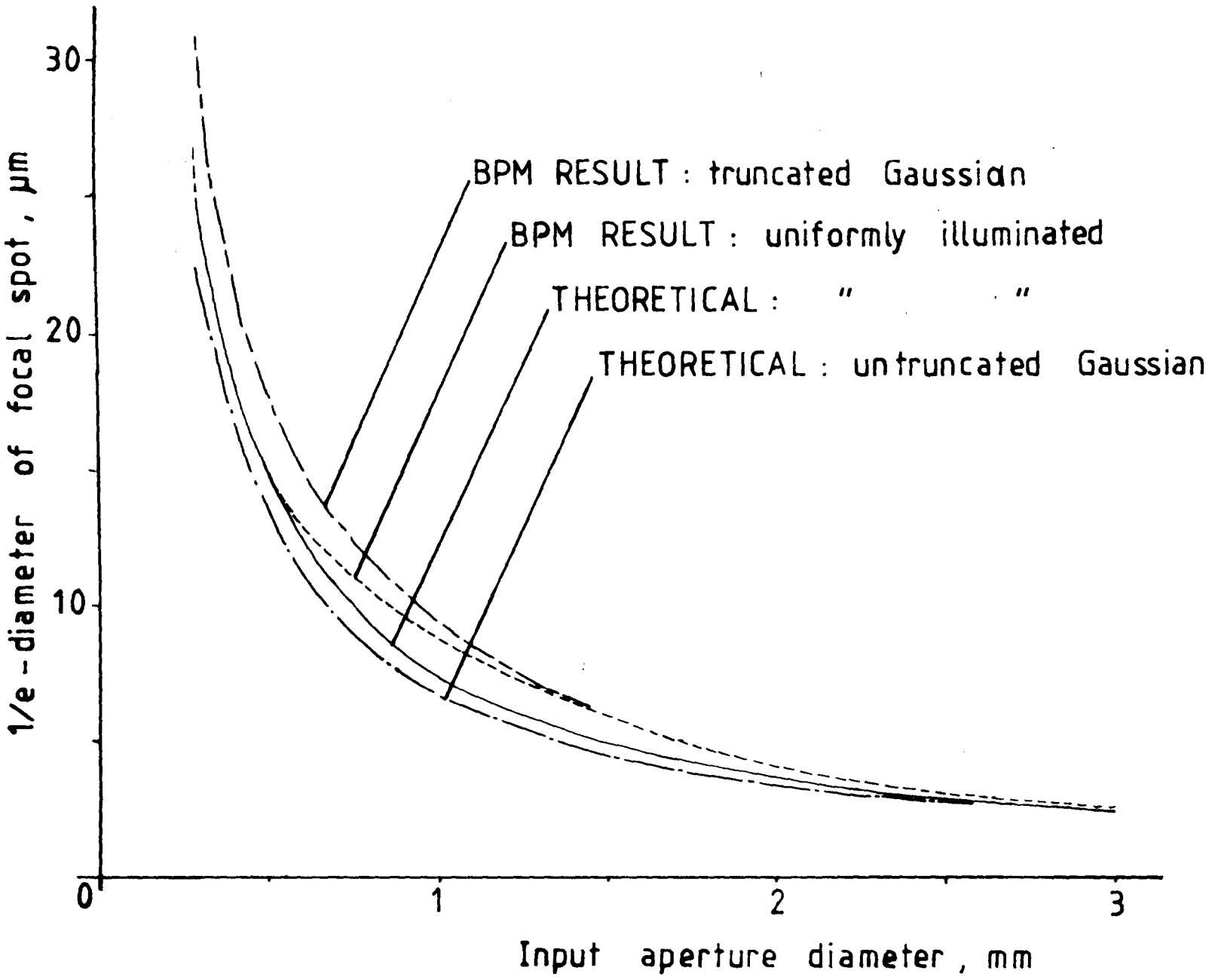


FIGURE [5,25] VARIATION OF FOCAL SPOT-WIDTH OF 'SOTTINI' GEODESIC LENS FOR DIFFERENT CONDITIONS OF ILLUMINATION

will generate a wider spot, since the uniformly illuminated aperture then carries more field energy at the edges of the aperture. Such behaviour is observed in Figure [5.25].

The final result of this chapter is given in Figure [5.26] where the evolution of an optical wave-field through the focus of the spherical geodesic lens with a rounded-edge specified in chapter two is shown. The incident field was a weakly truncated Gaussian, 1.0 mm in diameter and the field is focused at a distance of 50.23 mm from the centre of the lens, in comparison with the paraxial focus of 50.9 mm calculated in chapter four, a negative focal-shift thus also being exhibited in this case. The shift may be beneficial, as the lens suffers from geometrical aberrations which would tend to move the best focus well beyond the paraxial value (away from the lens).

5.1| Conclusions

On the evidence of focal fields obtained using the BPM under a variety of incident conditions, the BPM has been confirmed as a useful tool in the analysis of inhomogeneous lenses. Negative focal shifts have been calculated, using the BPM, which are in moderate agreement with published results. The shifts may be beneficial in lenses which suffer from positive aberrations, but would be unattractive for nominally diffraction-limited lenses, for, although diffraction-limited performance is very nearly attained in terms of focal spot-sizes and sidelobe levels, the a priori specification of the focal-length of a lens would be exceedingly difficult.

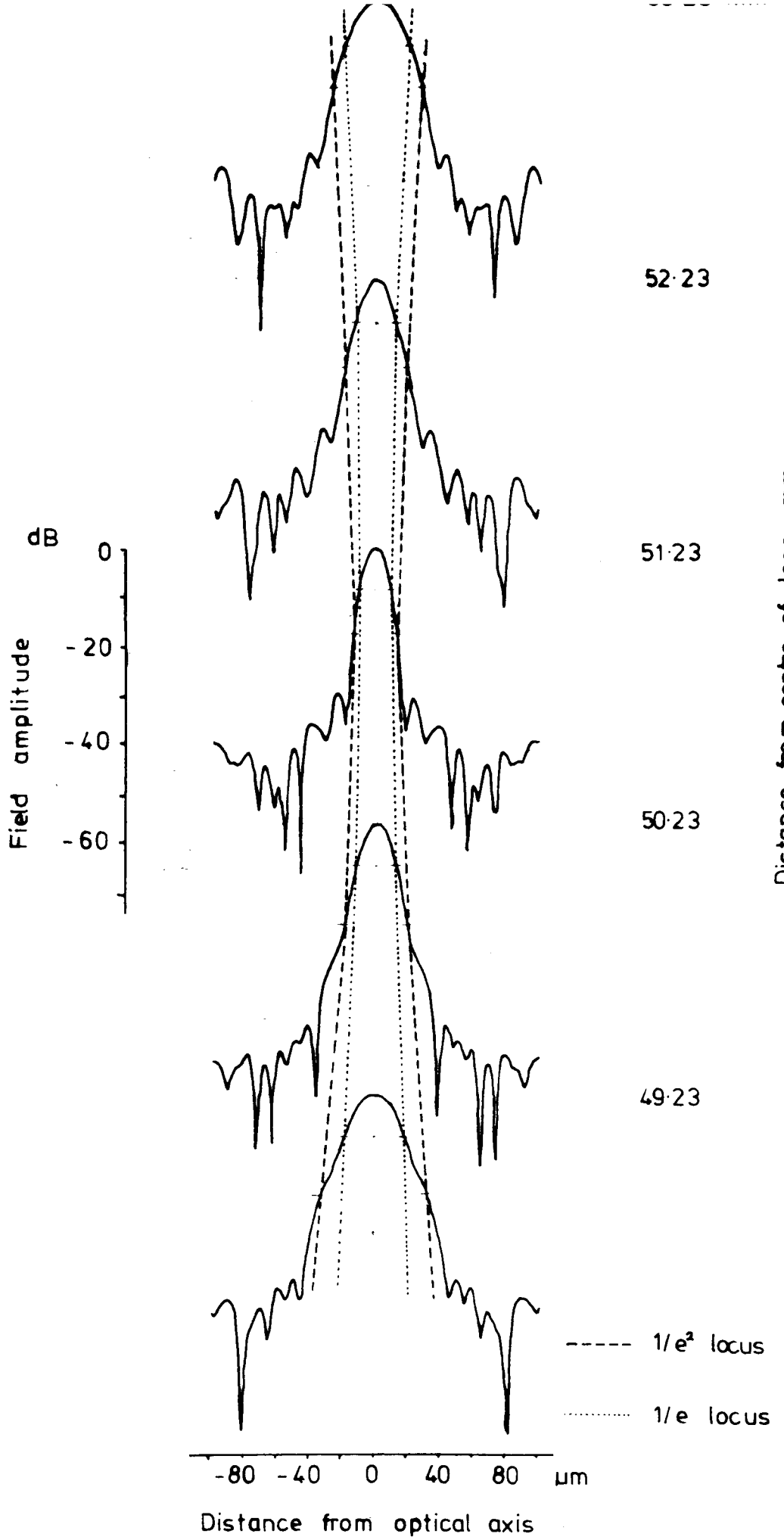


FIGURE [5,26]

field development near focal plane of lens SL 1

obtained using BPM

REFERENCES FOR CHAPTER FIVE

[1] Goodman,J.W.,
'Introduction to Fourier Optics', McGraw-Hill, New York, 1968

[2] Gaskell,J.D.,
'Linear Systems,Fourier Transforms, and Optics', Wiley and Sons, New York, 1978

[3] Brigham,E.O.,
'The Fast Fourier Transform', Prentice-Hall Inc., New Jersey, 1974

[4] Heideman,M.T., Johnson,D.H., Burrus,C.S.,
Gauss and the history of the Fast Fourier Transform, IEEE ASSP Magazine, Oct. 1984, 14-21

[5] Sziklas,E.A., Siegman,A.E.,
Diffraction calculations using Fast Fourier Transform methods, Proc. IEEE, 62(3), 1974, 410-412

[6] Van Roey,J., van der Donk,J., Lagasse,P.E.,
Beam-propagation method: analysis and assessment, J. Opt. Soc. Am., 71(7), 1981, 803-810

[7] Arfken,G.,
'Mathematical Methods for Physicists', 2nd ed., Academic Press, New York, 1970

[8] Feit,M.D., Fleck,J.A.,
Light propagation in graded-index optical fibers, Appl. Opt., 17(24), 1978, 3990-3998

[9] Fleck,J.A., Morris,J.R., Feit,M.D.,
Time-dependent propagation of high energy laser beams through the atmosphere, Appl. Phys., 10, 1976, 129-160

[10] Lagasse,P., Baets,R.,
The beam propagating method in integrated optics, NATO Advanced Research Workshop on Hybrid Formulation of Wave Propagation and Scattering, Aug 30-Sept 3, Rome 1983

[11] Kogelnik,H., Li,T.,
Laser beams and resonators, Proc. IEEE, 54(10), 1966, 1312-1328

[12] Born,M., Wolf,E.,
'Principles of Optics', Pergamon Press, Oxford, 1975

[13] van der Donk,J., Lagasse,P.E.,
Analysis of geodesic lenses by beam propagation method, Electron. Lett., 16(8), 1980, 292-294

[14] van der Donk,J., Vandewege,J.,
Second-order effects in geodesic lenses, Proceedings of the First European Conference on Integrated Optics, 14-15 Sept 1981, IEE

[15] Li,Y., Wolf,E.,
Focal shift in focused truncated Gaussian beams, Opt. Commun., 42(3), 1982, 151-156

[16] Li,Y,, Wolf,E.,
Focal shifts in diffracted converging spherical waves, Opt. Commun., 39(4), 1981, 211-215

[17] Saga,N., Tanaka,K., Fukumitsu,O.,
Diffraction of a Gaussian beam through a finite aperture lens and the resulting heterodyne efficiency, Appl. Opt., 20(16), 1981, 2827-2831

[18] Parker-Givens,M.,
Focal shifts in diffracted converging spherical waves, Opt. Commun., 41(3), 1982, 145-148

[19] Holmes,D.A., Korka,J.E., Avizonis,P.V.,
Parametric study of apertured focused Gaussian beams, Appl. Opt., 11(3), 1972, 565-574

CHAPTER SIX

FABRICATION

WAVEGUIDES AND INHOMOGENEOUS OVERLAY LENSES

6.1 Introduction

Experimental techniques used for fabricating waveguides and overlay lenses will be reported in this chapter. The substrate material used was lithium niobate, and ambient waveguides were formed using the technique of titanium in-diffusion. The material used for forming the lens was arsenic trisulphide, a high-refractive-index chalcogenide glass. The important properties of the substrate, the ambient waveguide and the overlay material will first be reviewed. The fabrication of planar waveguides and their observed optical properties will be discussed. Subsequent modifications of the fabrication environment to allow the formation of overlay lenses will then be discussed, and measured lens-profiles will be presented. Finally, an alternative method of producing overlay lenses will be suggested.

6.2 Materials

6.2.1 The substrate material

Lithium niobate, $LiNbO_3$, is a uniaxial ferroelectric crystalline insulator at room temperature with a number of interesting properties; in combination, these make it perhaps the most commonly used substrate material in integrated optics. It has also found widespread use in other fields, for similar reasons. The ferroelectric property means that the crystal possesses a net electric dipole moment which is reversible under an applied electric field. The ferroelectric characteristic is retained up to a very high temperature , 1140°C (T_c, the Curie temperature), which is only 115°C below the melting point of the crystal [67]. Above T_c, the crystal becomes paraelectric, ie no net charge then exists in the medium. [1]

Two grades of crystal are commonly available:

(a) electro-optic, or optical, high-purity grade which is nearly single domain and is optically highly homogeneous. This is the most difficult and therefore most expensive type to manufacture.

(b) acoustic, or transducer grade which is of lower purity and can have micro-domains at the surface but which is required to be of good quality in other respects.

The properties and characteristics of lithium niobate which make it of importance in modern optics are:

(a) a large negative birefringence at visible to infra-red wavelengths from 0.4-5.0 um [2];

(b) fairly large absolute values of refractive index [2];

(c) good transmission properties in the visible and infra-red regions of the spectrum [2];

(d) a large non-linear optical coefficient, making the material an excellent one for non-linear optical investigations such as phase-matched second-harmonic generation (SHG), sum- and difference-frequency generation and parametric effects [2,3]. The phenomenon of optical damage in the material (see below) is a drawback, however,since optical power levels must be restricted;

(e) strong piezoelectric surface-wave properties [4] which renders the material a crucial choice in many surface acoustic wave (SAW) applications;

(f) excellent acoustic properties, particularly with regard to propagation loss, at frequencies up to and beyond 1GHz [5,6];

(g) good acousto-optic properties. White et al [7] have shown that this is largely due to an indirect effect in which an applied acoustic stress in the piezoelectric crystal generates an electric field that in turn causes a change in refractive index through the electro-optic effect. The direct photoelastic contribution is relatively small;

(g) high electro-optic coefficients together with high electrical resistivity which allow for easy application of and good interaction with modulating fields, making the crystal excellent for electro-optic devices [8].

For these reasons, lithium niobate was the substrate material chosen for most of the investigations reported herein.

A note of caution should be sounded in order to redress what otherwise seems to be a most favourable balance, for the material has some deficiencies. The crystal has been observed to suffer index inhomogeneities (optical damage) at fairly low values of optical power [9] . This effect is due to the photo-excitation of electrons in the material which then drift under the influence of fields internal to the crystal. The drifting causes local index inhomogeneities through the electro-optic effect. The extra-ordinary index n_e is most stongly affected with a decrease, albeit reversible, occurring. The effect of optical damage can be minimised by operating at longer wavelengths. A further problem is the crystal anisotropy. In some cases, eg mode conversion experiments in acousto-optics [10], anisotropy presents a distinctly favourable aspect. However for low f-number lens applications anisotropy introduces polarization effects that can require for example that vector field theories be used in

analysis [11]. Van der Donk [12] found that crystal anisotropy was not likely to cause significant effects in lenses of moderate f-number fabricated on $LiNbO_3$, and as such anisotropic effects were not a major cause of concern in the present study. Vahey et al have indicated that anisotropic effects could be important [13] in geodesic lenses with very low f-numbers, especially for waveguide modes that are well-confined, ie far from cut-off.

6.2.2 The ambient waveguide

The technique of establishing a high-index waveguiding region on the surface of a lithium niobate substrate by means of titanium metal indiffusion has been known for over a decade [14,15]. Other metals such as nickel, vanadium [14], manganese, zinc, copper, cobalt, and chromium [16] have been tried but titanium has generally produced the best results. A thin (200-1000A) film of titanium is evaporated onto the crystal surface under vacuum conditions and is then diffused into the crystal by heating in a high temperature furnace. Conditions of fabrication can vary quite considerably [15], but to a first approximation the maximum achievable index change depends only on d, the evaporated film thickness, whilst the effective guide depth is controlled by the diffusion time t and the diffusion temperature T (which is almost invariably less than the Curie temperature) [14]. In consequence the most important parameters of the waveguide are controlled independently, in contrast to the situation that exists for another well known technique of fabricating a waveguide on lithium niobate, that of out-diffusing lithium ions by heating the virgin crystal at elevated temperatures [17].

The titanium concentration profiles, which determine the refractive index profiles, resemble complementary-error functions for short diffusion times and Gaussian functions [18,16] for longer diffusion times. Commonly used diffusion times are those for which the metal is just completely in-diffused, or slightly longer. Much research is still going on into the exact processes involved in waveguide formation. Other fabrication conditions

which have been found to bear upon the properties of the waveguide are:

(a) the ambient gas(es) present during the diffusion cycle,

(b) the presence or absence of lithium niobate powder in the diffusion chamber. The powder is thought to suppress unwanted out-diffusion [19],

and (c) the re-oxidation conditions.

The precise conditions used in the present study will be reported in a later section. The method produced good planar guides possessed of low loss and in- and out-of-plane scatter.

The question of obtaining waveguide uniformity will be discussed later in this chapter, and in chapter eight. It is extremely difficult to fabricate a uniform waveguide on geodesic lenses using vacuum deposition. Alternative techniques will be pointed out which may well obtain greater uniformity though at the expense of poorer waveguiding; a trade-off has then to be established.

Ti:$LiNbO_3$ has been used extensively in integrated optics for the formation of devices such as directional couplers [20], switches [21], high bandwidth amplitude and phase modulators of both the lumped element [22] and travelling-wave [23] types, gratings [24], fibre- to integrated-optic couplers [25], geodesic lenses [26,27], acousto-optic modulators and deflectors [10], integrated optical spectrum analyzers [28,29] and many others.

6.2.3 Material used for the inhomogeneous overlay lens

Amorphous arsenic trisulphide (As_2S_3) is a non-oxide chalcogenide which forms one of the most stable of the sulphide glasses [30]. It is thought to consist of a two-dimensional network structure with some short-range ordering [31] in the bulk

state. The short-range ordering is essentially the same as in crystalline As_2S_3.

Films of arsenic trisulphide can be produced by rf sputtering [32] or by thermal evaporation from a resistance-heated boat [33]. The properties of sputtered films conform closely to the bulk structure, but the as-evaporated films do not [34]; for example, the density, characteristic acoustic impedance and the refractive index are 6-13% lower than the bulk values [35]. If the films are subsequently annealed, however, the properties once more approach the bulk vitreous state.

Bulk samples are reddish-brown in colour, whereas evaporated films fall somewhere between pale yellow and orange, depending on the film thickness. Arsenic trisulphide is almost insoluble in water and exhibits a very low hygroscopicity (water absorption) [30], clearly a very desirable characteristic in a thin film optical material [36]. The material is soluble however in even mild alkaline solution and this can be used to fabricate patterns. Small amounts of hydrogen sulphide are given off in reaction with alkali. Because of the low solubility in water arsenic trisulphide is not thought to be acutely poisonous. However, like other arsenious compounds, it is easily absorbed through the skin and mucous membranes. Furthermore, under conditions of elevated temperatures such as can occur in thermal evaporation chambers, noxious compounds can form and in consequence reasonable caution has to be exercised when dealing with the material.

The glass transition temperature, T_g is 470^oK [37]. The melting point of the glass is 483^oK [34]. Unannealed films are easily scratched and can be wiped off the underlying substrate by rubbing; annealed films are more resilient but must still be treated with care.

One of the most interesting properties of arsenic trisulphide (a property shared with some other chalcogenides) will now be dealt with. The study by De Neufville, Moss and

Ovshinsky on photo-structural transformations in amorphous chalcogenide films [34] was extremely important. Each of the main conclusions of the study triggered off new and useful discoveries in optical engineering processing and components. The important optical characteristics of evaporated As_2S_3 films listed by de Neufville et al were to have significant consequences for the present study.

1) Slowly evaporated amorphous films of arsenic trisulphide differ from the bulk material as mentioned above. If exposed to radiation from a white-light (eg tungsten source), the films undergo a structural transformation which may be termed photo-structural;

2) a similar structural (thermo-structural) transformation takes place if the evaporated film is annealed at a temperature approximately equal to the glass transition temperature, T_g;

3) both of the above changes lead to an irreversible increase in the refractive index of the films, which attain values close to that of the bulk material. This optical transformation is thought, in both cases, to be associated with a polymerization of the molecular units as initially deposited, itself an irreversible structural change;

4) illumination of thermally-annealed films or, conversely, annealing of illuminated films leads to no further measurable change in refractive index;

5) the evaporated films, once exposed to either annealing or illumination, possess optical and structural properties close to both the bulk samples and to sputtered films (sputtered films exhibit only weak photo- or thermo- structural effects);

6) annealed films, if illuminated in the ultra-violet, exhibit a decrease in the absorption edge energy of approximately

0.03eV which can be reversed (bleached) by re-annealing. This reversible effect is called photo-darkening.

Two modes of photo-response were distingushed, therefore, by De Neufville et al in evaporated arsenic trisulphide films (the study also concerned itself with films of arsenic triselenide). One, termed photo-structural, is characterized by changes in both the position of the optical absorption edge and the value of refractive index upon illumination. The other photo-effect, termed photo-darkening, is characterized by a reversible absorption edge shift.

De Neufville et al also discovered that the chemical reactivity of As_2S_3 was enhanced after exposure to illumination. They observed that exposed films were invariably oxidised even in vacuums of less than 10^{-6} Torr, whilst unexposed films were not. In a later study, Suhara et al found that exposed As_2S_3 was preferentially etched by the action of mild NaOH solution 1.79 times faster than the unexposed film. This effect was utilised by them to form waveguide holograms with a diffraction efficiency of 92% [38].

The thermally erasable photo-darkening effect has also been used to construct waveguide holograms, with a diffraction efficiency of up to 80% and good homogeneity of phase-relief [39]. Such a process is characterized by the absence of a development stage. As_2S_3 and other chalcogenide materials are consequently of increasing importance in the field of optical storage.

Other developments incorporating As_2S_3 in integrated optics include the fabrication of planar and strip waveguides [32,40], electro-optic modulators [41], a thin-film acousto-optic deflector possessing a deflection efficiency of 93% [42] (the acousto-optic figure-of-merit M2 of As_2S_3 is among the highest of all materials [43]), an acousto-optic convolver [44], switching devices [45], fiber-waveguide taper couplers [33], graded index waveguide Fresnel lenses [46] and micro-gratings [47]. The latter

two references are particularily interesting, as a further means of inducing photo-sensitivity was demonstrated. Electron-beam irradiation was used to cause refractive index changes of up to 4%. Such a procedure is exciting because of the high resolution and placement accuracy that can be achieved using computer control. The ultimate line resolution of typical chalcogenide films is expected to be high, of the order of 10,000 lines/mm.

A possible problem associated with the formation of thin-films of As_2S_3 on $LiNbO_3$ substrates is that the thermal expansion coefficients of the two materials are different. $LiNbO_3$ has anisotropic thermal expansion coefficients $\alpha_a = 16.7 \times 10^{-6}/^{o}K$ and $\alpha_c = 2 \times 10^{-6}/^{o}K$ [53] whereas As_2S_3 has a thermal expansion coefficient of $24.6 \times 10^{-6}/^{o}K$ [37]. The fact that As_2S_3 has a higher value is undesirable since the result is that the film is constantly under tension rather than compression. Tension can lead to cracks and structural weaknesses, and may therefore be a factor contributing to the high in-plane scattering levels that are found in these waveguides.

The properties of As_2S_3 that influenced its use in the work reported herein were: a high refractive index relative to $LiNbO_3$; high optical transmission at longer visible and infra-red wavelengths; easy fabrication of thin-films; and the possibility of tuning the focal characteristics using the photo-sensitive effects described above.

6.3 Fabrication of planar waveguides

6.3.1 Fabrication of Ti-diffused $LiNbO_3$ waveguides

Acoustic-grade $LiNbO_3$ was used as the substrate for most planar waveguide experiments. High-quality substrates were used for lens fabrication. Y-cut $LiNbO_3$ was almost invariably used, with optical propagation intended to be in the crystal x-direction. The rationale behind this choice was that, in applications where the lenses were to be used as spatial Fourier transforming elements with the input signal presented as a

surface acoustic wave, the highest achievable signal fractional bandwidths would be obtained with the acoustic wave travelling in the z-direction on a y-cut crystal [48].

Crystal plates as obtained from the manufacturer were 50mm in diameter and 0.5mm or 1.0mm thick. The crystal z-direction was indicated by a flat. The plates were optically polished on one y-face and roughly ground on the other to minimise unwanted reflections. The plates were sawn into 8mm(z-direction) x 20mm-40mm (x-direction) substrates which were then checked for cracks and defects under 40x microscope magnification. Those that exhibited defects were rejected. Doughty [48], in his work on the fabrication of the geodesic lenses that were optically characterised by the present author, prepared several high-quality $LiNbO_3$ substrates. The procedures used by him produced substrates which were defect-free under 100x magnification, flat to within 2 um, and plane-parallel to within 20 seconds of arc. These substrates were all drawn from one high-purity boule and had dimensions of 55mm x 14mm x 2.95mm, in order to support two geodesic lenses of diameter 10mm and focal length 18.5mm. Several of the substrates were made available to the present author to allow fabrication of inhomogeneous Luneburg lenses made from As_2S_3, together with ambient waveguides.

The substrates were cleaned prior to waveguide formation using a procedure standard to the Clean Room of the Department of Electronics and Engineering at the University of Glasgow. The procedure has been proved adequate to the preparation of substrates for the fabrication of stripe waveguides, a considerably more demanding task than that of producing planar waveguides. The procedure is as follows:

1. Immerse sample in trichloroethylene and agitate ultrasonically for 10 minutes.

2. Rinse off trichloroethylene by immersing sample in methanol. Agitate ultrasonically for 1 minute.

3. Immerse sample in clean methanol and agitate ultrasonically for 10 minutes.

4. Immerse sample in acetone and agitate ultrasonically for 10 minutes.

5. Immerse sample in 5% DECON 90 solution and agitate ultrasonically for 10 minutes.

6. Immerse sample in acetone and agitate ultrasonically for 20 minutes.

7. Immerse sample in methanol and agitate ultrasonically for 5 minutes.

The samples were not allowed to dry between any of the above stages as it has been found that contamination can often take place as a result. After stage 7 the samples were rinsed in distilled water, blown dry and placed in the high-vacuum evaporation system used for depositing titanium metal films. Titanium was then evaporated using an electron gun source and deposited onto the substrate, at a vacuum of 1×10^{-5} Torr. The deposited film thickness was monitored using a calibrated quartz crystal monitor. Typical thicknesses required to produce a single mode waveguide were 18.0nm to 25.0nm, depending on subsequent diffusion conditions. Some substrates had thicker films deposited on them, in order to allow multi-moded waveguides to be studied. The deposited metal film thickness was checked after deposition by a Talystep stylus instrument.

The procedure for diffusing the titanium metal into the $LiNbO_3$ substrate has varied considerably among different groups of workers [49,50,19,14]. The degree of in-plane scattering was of particular importance in the present study since scattering degrades the focal intensity pattern produced by a lens. Scattering has several causes, notably the surface roughness at the waveguide interfaces, especially the air/guide interface; defects or impurities associated with the bulk crystal; and

defects introduced during, or associated with, the waveguide fabrication process. The first two problems can be minimised by polishing and selecting high-purity material respectively. The latter problem was dealt with in this study by investigating three reported techniques and selecting the most favourable one.

Three techniques which have been reported are:

(i) diffusion of the Ti layer in the presence of congruent $LiNbO_3$ powder in a closed tube [19]. This technique is thought to suppress the out-diffusion of Li_2O which can take place at elevated temperatures. Out-diffused Li_2O can cause parasitical waveguiding effects which interfere with the operation of stripe guides, for example;

(ii) diffusion of the Ti layer in a flowing atmosphere of inert gas (usually argon) in an attempt to prevent oxidation of the Ti layer. Cooling subsequently takes place in oxygen gas in order to re-oxidise the $LiNbO_3$ [25]. This technique is thought to reduce susceptibility to optical damage;

(iii) diffusion of the Ti layer in a flowing air or oxygen atmosphere, in order to inhibit the precipitation of $LiNb_3O_8$ which can act as a local scattering centre [49]. The formation of $LiNb_3O_8$ is inhibited when indiffusion is performed in a wet atmosphere [56].

Application of the first technique typically involved a ten hour diffusion time at 1000°C. The second involved a nine hour diffusion in a flowing, wet argon atmosphere at 1000°C, followed by flushing for 2 minutes in dry argon and finally a further one hour diffusion in dry oxygen. The third technique was similar to the second except that wet air or oxygen was used throughout the diffusion.

Results obtained under nominally identical conditions, in terms of initial Ti thickness, diffusion time and temperature, consistently showed that the first technique was the most

unfavourable. Prism-coupled m-lines using this technique displayed considerable scattering with poor angular confinement of the propagating modes. The other two techniques displayed good confinement and had little to choose between them. The inert gas technique was chosen for convenience as it had found favour with other workers in the department. Figure [6.1] shows a typical m-line intensity scan obtained in the far-field from a waveguide fabricated using the inert gas technique. The light level at 1^{o} relative to the peak is -31dB. Vahey [51] has reported -40dB at 1^{o}.

No attempt was made to measure the attenuation of these waveguides though no more than 2dB/cm was expected.

6.3.2 Arsenic Trisulphide Waveguides

6.3.2.1 Fabrication

The fabrication system and conditions used in the formation of thin-film planar waveguides of As_2S_3 in the present study were those described by Stewart et al [33]. After preparation of the substrates in the fashion described in section 6.3.1, films of As_2S_3 were thermally evaporated from a tantalum enclosed-crucible source. The bulk material had been stored in dark conditions to prevent any unwanted photostructural effects. The films were thermally annealed after deposition. The values of important parameters in the deposition and annealing processes are summarised in Table [6.1]. An optical thickness monitor was used to control the deposited film thickness to an accuracy of 10 nm. The monitor was also capable of detecting changes in film homogeneity during deposition. The operation of the monitor is described in detail in reference [33].

6.3.2.2 Evaluation

After annealing in a dry nitrogen atmosphere for 50 minutes, the system was allowed to cool for 24 hours before the samples were removed. Annealed films were found to be much stronger and

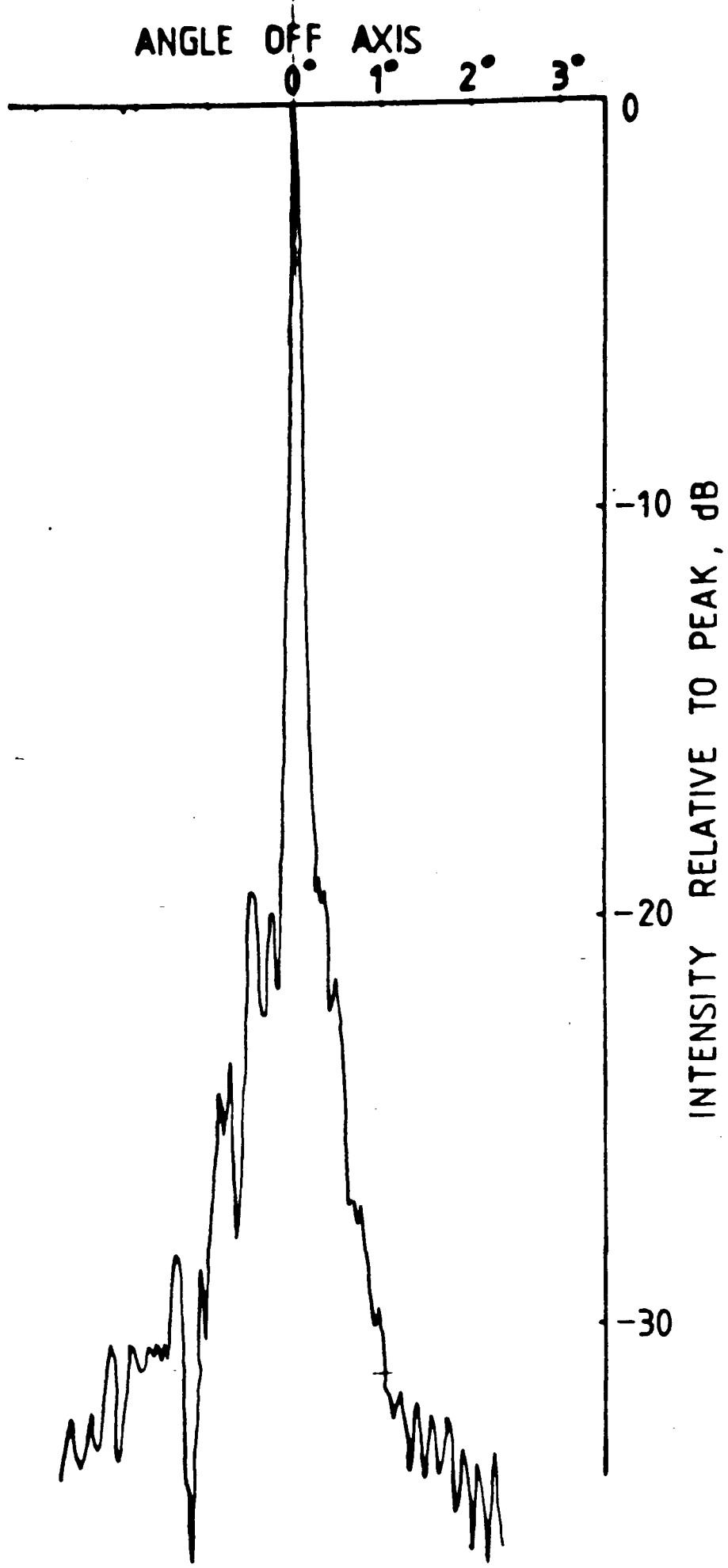


FIGURE [6.1] m_o-line intensity, far-field scan
Ti: $LiNbO_3$ - method (ii)

Evaporation temp. (°C)	312 ± 1
Vacuum pressure ($\times 10^{-5}$ Torr)	< 1·0
Typ. deposition rate (Å/sec)	30·0
Substrate temperature during deposition (°C)	25·0 typ. 150·0 max.
Annealing temp. (°C)	200·0
" time (mins)	50
" conditions	in Nitrogen atmosphere, at atmospheric pressure.
Typical film index	
(a) before anneal	2·444
(b) after anneal	2·6

θ λ = 633nm

TABLE [6.1] As_2S_3 WAVEGUIDE FABRICATION CONDITIONS

harder than un-annealed films, which could be wiped off the substrates with vigorous rubbing. Scotch-tape adhesion tests revealed however that even the annealed films could be peeled off both glass and $LiNbO_3$ substrates.

The thickness of deposited films was measured, post-fabrication, using Talystep, ellipsometry and waveguiding techniques. The latter two methods were also used to measure the refractive indices of the films, although ellipsometry techniques were not particularily sensitive in this respect. The thickness measurements showed good correspondence with each other. The refractive indices of the annealed films averaged 2.596 at a wavelength of 633nm, with a small standard deviation of +/- 0.008 .

A decrease in film thickness was found to occur as a result of the annealing process; for example, a film measuring 574nm before annealing was found to be 556nm thick after annealing. Assuming that such a change is purely due to an increase in film density rather than a result of partial re-evaporation, the corresponding increase in film density is 3.2% .

6.3.2.3 Optical evaluation

As_2S_3 waveguides produced considerably higher levels of scattered light than did Ti:$LiNbO_3$ waveguides. This is evident from Figure [6.2] which shows contour plots of the scattered light intensity in the far-field for m-lines coupled out from both types of waveguide using a prism. The mode-line from the Ti-diffused waveguide is well-confined, whereas the mode-line(s) from the As_2S_3 waveguide are heavily scattered. Furthermore, parasitical coupling of energy into other modes of the waveguide is evident.

A typical far-field scan of a TE_0 mode obtained from a 920nm thick waveguide at a wavelength of 633nm is shown in Figure [6.3]. The amount of scattered light is clearly considerable. At 1^o off-axis the intensity level with respect to the peak value is

-10.7dB . At 2^{o} a local peak displays an intensity level of -8.6dB, and at 3^{o} the intensity is still only -16.1dB. Further evidence of high scatter levels in As_2S_3 waveguides is presented in Figure [6.4]. These photographs show the prism-coupled far-field mode lines observed for a film 743nm thick, at two wavelengths. Photographs (a) and (b) of the figure were taken at 633nm, whereas photographs (c) and (d) were taken at 1150nm. The scattering levels are very high in (a) and (b) and, furthermore, a great deal of energy is being transferred to modes other than the one being excited at the input prism-coupler. At 1150nm photographs (c) and (d) show that the level of scattering is much reduced. No coupling of energy into parasitical modes was observed either, though the film was capable of supporting two modes at this wavelength. Improved performance of these waveguides at the longer wavelength was consistently observed. The length of the 'streak' arising from out-of-plane scattered light in the waveguide was rarely longer than 2.0cm at 633nm and fell much below 1.0cm for very thin waveguides. Using Tien's method of estimating waveguide attenuation according to the sensitivity of the eye [54], such figures indicate losses of between 13dB/cm and 27dB/cm. Measurements made using a Hamamatsu TV intensity-display system under magnification indicated that losses of greater than 45dB/cm occurred for unannealed films. At longer wavelengths the scatter streaks lengthened considerably and rarely failed to propagate right to the end of the waveguide. The best loss measurement obtained at a wavelength of 633nm was for a 1200nm thick film deposited onto a glass waveguide. The peak intensity variation as a function of the distance separating two prisms on the waveguide is shown in Figure [6.5]. A total loss of 7.55dB/cm is observed.

Photo-refractive damage occurred at a wavelength of 633nm. Figures [6.6] and [6.7] show the effects of optical damage on the TE_0 mode, imaged onto a Hamamatsu camera and image-acquisition system. A three-to-four minute interval separated the acquisition of the two images. Before the second image could be acquired, the input prism-coupling angle had to be re-adjusted to optimise coupling of the mode, as indicated by the shift of the peak. Such

(a)

(b)

FIGURE [6,2] Contour plots of scattered light intensity in output m-lines

(a) Ti - diffused waveguide

(b) As_2S_3 thin-film waveguide

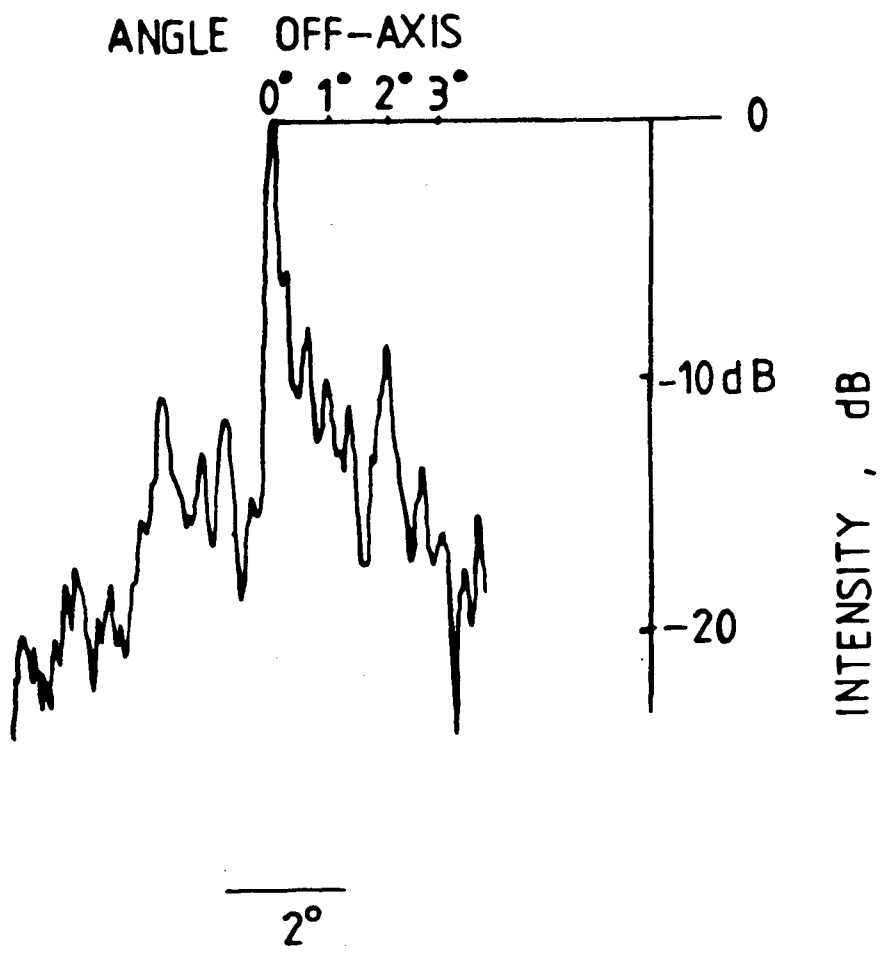
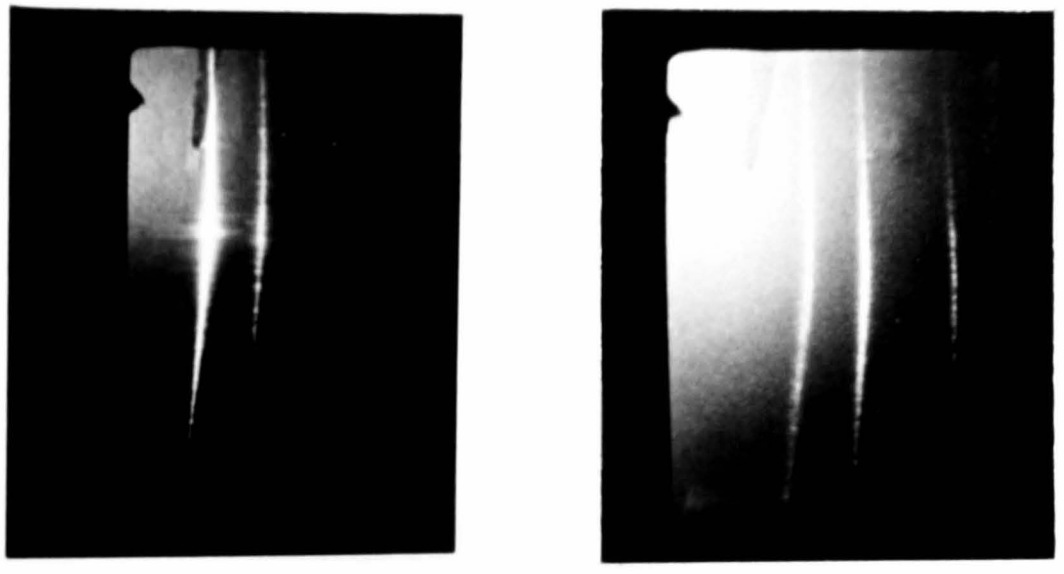
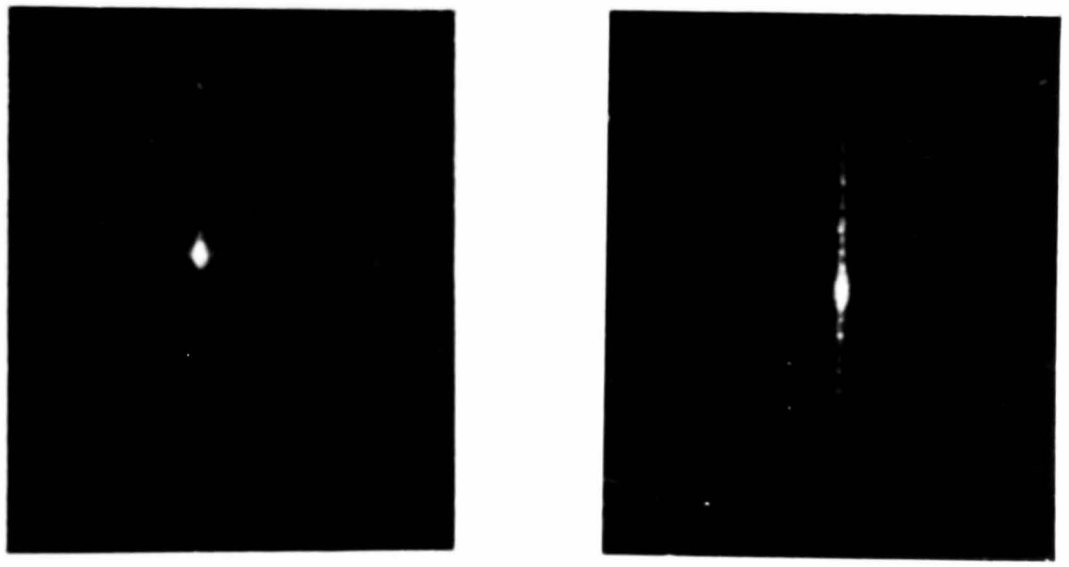


FIGURE [6,3] m_o - line intensity, far-field scan

As_2S_3 waveguide on y-cut $LiNbO_3$



(a) m_o $\lambda = 0{\cdot}633\mu m$ (b) m_1



(c) m_o $\lambda = 1{\cdot}15\mu m$ (d) m_1

FIGURE [6,4] Output m-lines observed using prism-coupling :- As_2S_3 waveguide on $LiNbO_3$

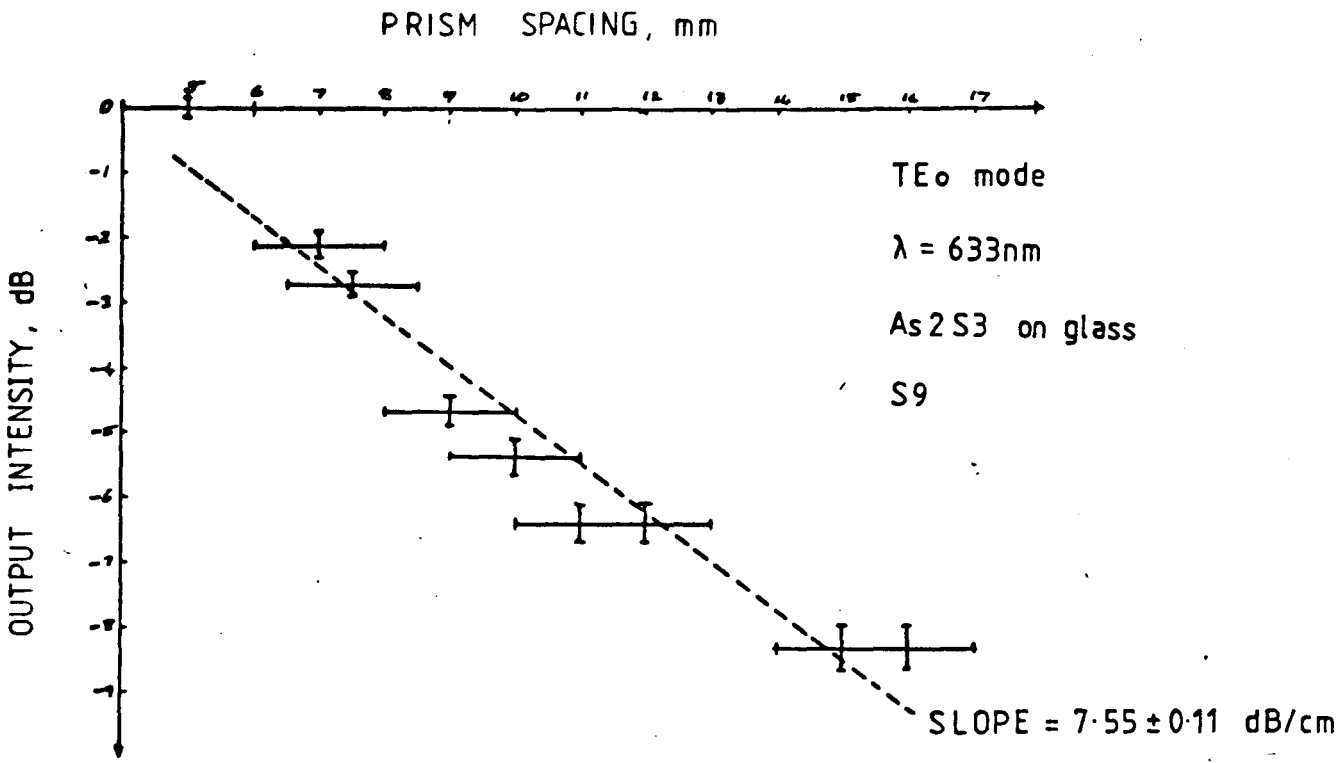


FIG [6,5] output intensity vs. prism spacing

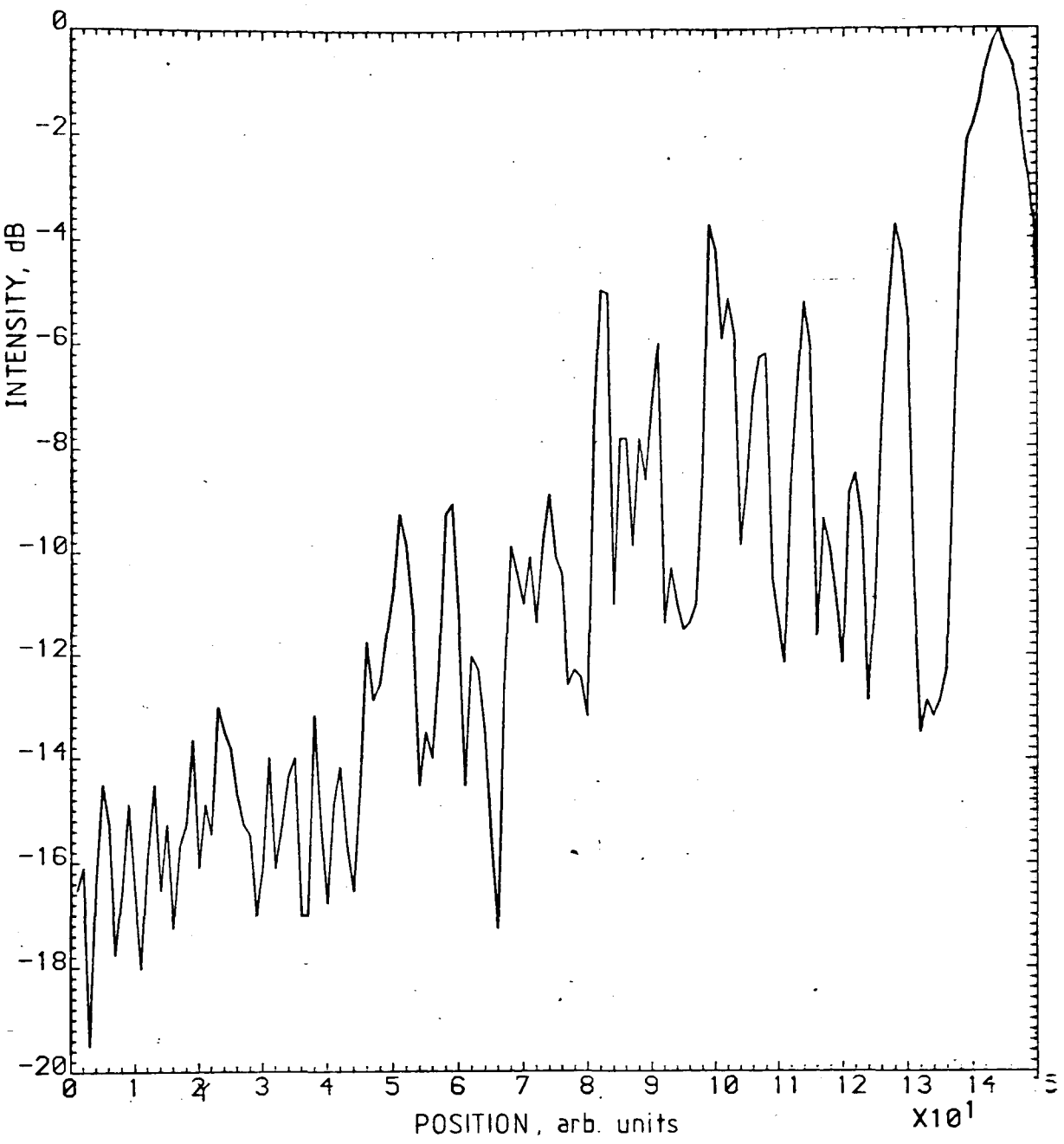


FIG [6,6]As2S3 Waveguide on glass λ = 633nm

TEo mode

high in-plane scatter, peak at edge of plot

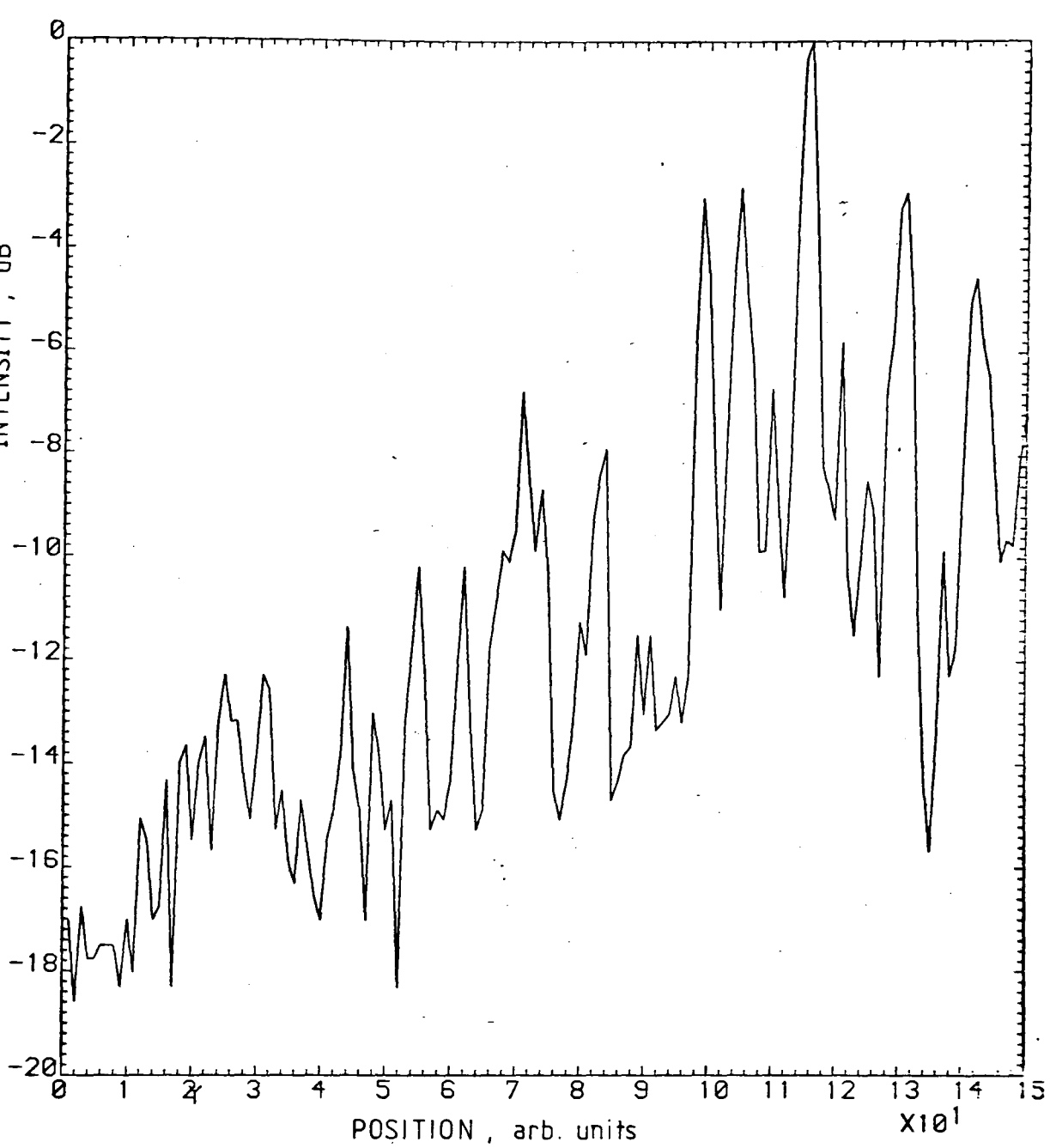


FIG [6,7] as fig [6,6] but 3-4 mins later;

peak has SHIFTED, NARROWED

and BROKEN UP

due to OPTICAL DAMAGE

a re-adjustment indicated that the refractive index of the waveguide had increased.

Films thicker than approximately 2500nm showed signs of 'blooming'- the films took on a milky grey hue- indicating that homogeneous growth is impossible at these thicknesses. Such films were not capable of guiding light waves.

6.4 Fabrication of inhomogeneous overlay lenses

6.4.1 Deposition of lenses

Two masks were used to fabricate inhomogeneous overlay lenses. One mask, designated the profile mask, was intended to shape the deposited thin-film to a profile corresponding as closely as possible to the Luneburg lens design profile desired. The Luneburg lens profiles were given in chapter two. The theory governing the aperture and placement of the profile mask was presented in chapter three. A second mask, designated the edge-definition mask, was intended to make fine adjustments to the edge-profile of the deposited lens. It was to be situated adjacent to the substrate and had an aperture diameter approximately equal to the desired lens diameter. Several masks of different aperture diameters were constructed to meet the design criteria for different lenses.

A photograph of the shadow-masking system is shown in Figure [6.8]. The support jig and masks were both of stainless steel construction. One leg of the support jig was gradated and threaded in order to facilitate mask placement. The masks were secured by screws to movable supports. These supports were secured to the threaded leg by two ring-bolts. The mask positioning accuracy was estimated to be approximately 0.5mm . Both profile and edge-definition masks were tapered to ensure that they acted as thin masks of the type investigated in chapter three.

The geometry of the experimental evaporation system is shown

FIGURE [6,8] Masking jig for fabricating inhomogeneous overlay lenses

in Figure [6.9]. A tantalum crucible (of the enclosed type, to prevent spitting) contained the As_2S_3 powder. The As_2S_3 powder was stored in side-compartments of the crucible, to prevent the target substrate directly 'seeing' the source. A shutter situated just above the source was used to expose the substrates to the molecular flux. After the films were deposited the test substrates were transferred to a heater for thermal annealing.

The thickness control and monitoring system used in the arrangement of Figure [6.9] was far from ideal. An optical thickness monitor had been used to monitor the fabrication of planar waveguides at a constant source-to-substrate distance of 200.0mm . It was found impossible to utilize the optical monitor in the fabrication of overlay lenses, since the shadow masks prevented the optical beam from impinging upon the target substrate. Furthermore, the wide spread of experimental conditions under which the lenses were subsequently to be fabricated would have necessitated considerable effort in calibrating the system. A quartz crystal thickness monitor was used instead. The profile mask had two additional openings one of which permitted the crystal thickness monitor to 'see' the As_2S_3 source, and another which supported a cover slip that was later used as a check on the deposited film thickness. Both crystal and cover-slip were calibrated against samples lying on the edge-definition mask. The edge of a second cover-slip was used to define a sharp edge on the sample that could later be measured with a Talystep.

An alternative method of controlling the film thickness was to determine the rate of deposition under controlled deposition conditions. Considerable information on the deposition rate was available from the experiments carried out on planar waveguides.

6.4.2 Measurement of fabricated profiles

A Talystep instrument adapted to operate in a long-range profile scanning mode was used to measure the profiles of as-fabricated lenses. The Talystep had a maximum traverse of 2.5mm

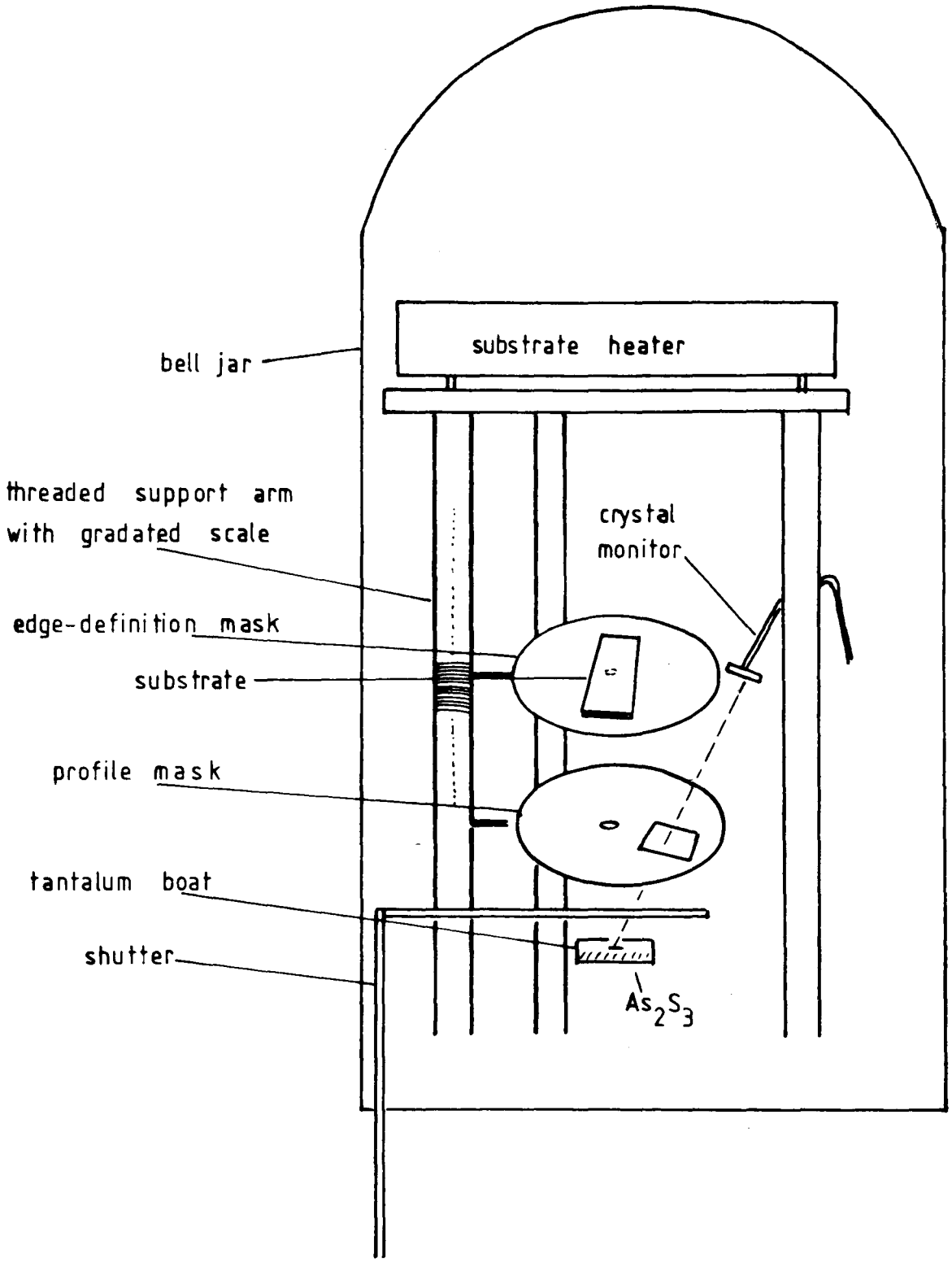


FIGURE [6,9] Geometry of the experimental evaporation system

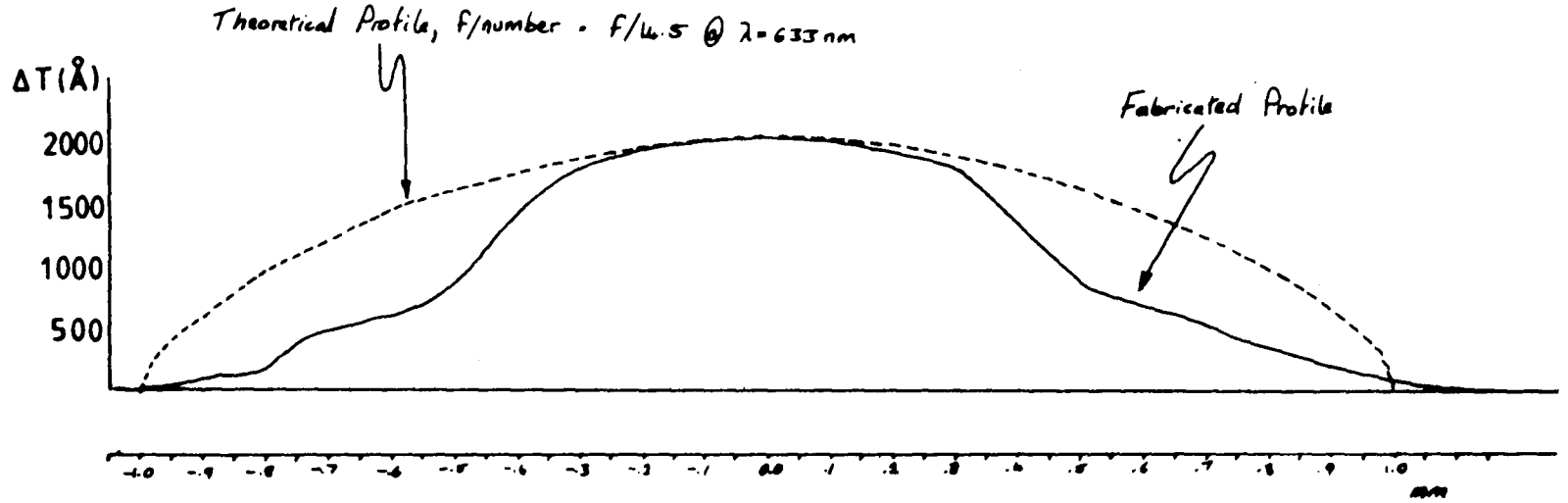
only, and was therefore of limited use in measuring entire profiles of large diameter lenses. It was mainly used to scan the edge-profiles of the lenses.

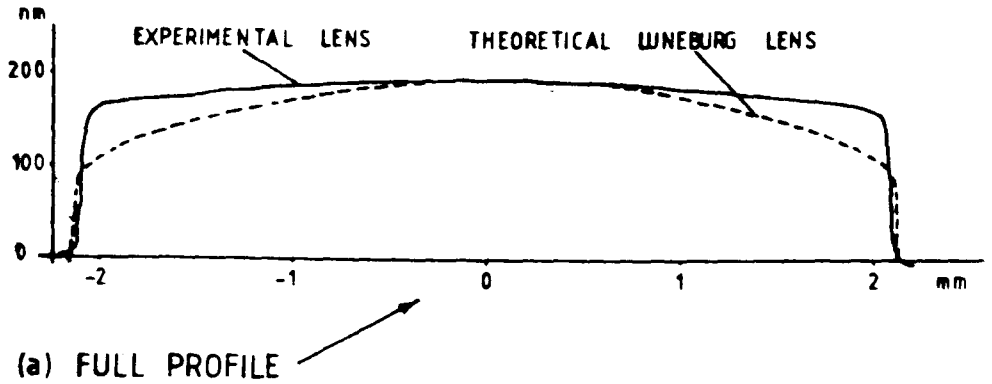
Figure [6.10] shows the experimental thickness profile of a relatively weak lens having a diameter of 2.0mm and a paraxial focal length of 29mm. The corresponding theoretical profile of the perfect Luneburg lens is also shown. The experimental lens was fabricated on a glass substrate, with the lens and waveguide both made from As_2S_3. It was shown in chapter two that the use of the same material for waveguide and lens could lead to considerable gains on the tolerances required to produce a lens of a given f-number. Lenses with long focal lengths could then be fabricated. The profiles were well matched in the central region. The overall profile was not unlike the profiles predicted in chapter three, except for the taper at the edge. An edge-definition mask had not been introduced into the system at this stage.

Lenses produced using the $Ti:LiNbO_3$ - As_2S_3 combination were much more difficult to fabricate due to the tolerances involved. As was shown in chapters two and three only lenses of small f-number were feasible and these required small source-to-substrate and source-to-mask distances. The problems involved in thickness control then became considerable. It was decided to fabricate some lens profiles on a largely empirical, trial-and-error basis, at longer distances than those prescribed by the shadow-masking theory. Figure [6.11] (a) and (b) show a complete lens profile and edge scan obtained under such conditions. The profile was obtained using a stylus-operated computer-controlled Tencor profileometer which had a maximum traverse of 6.5mm. The profile theoretically required is also shown in the Figure. A well-controlled centre thickness of 1850nm was obtained with this lens due to fabrication at a source-to-substrate distance of 50mm, with the shadow-mask situated at 12.0mm form the source. The overall profile of the lens fabricated, however, was flatter than required, as expected from the distances involved.

FIG [6,10] Lens profiles

As_2S_3 lens / As_2S_3 waveguide / glass substrate





(a) FULL PROFILE

(b) EDGE OF PROFILE

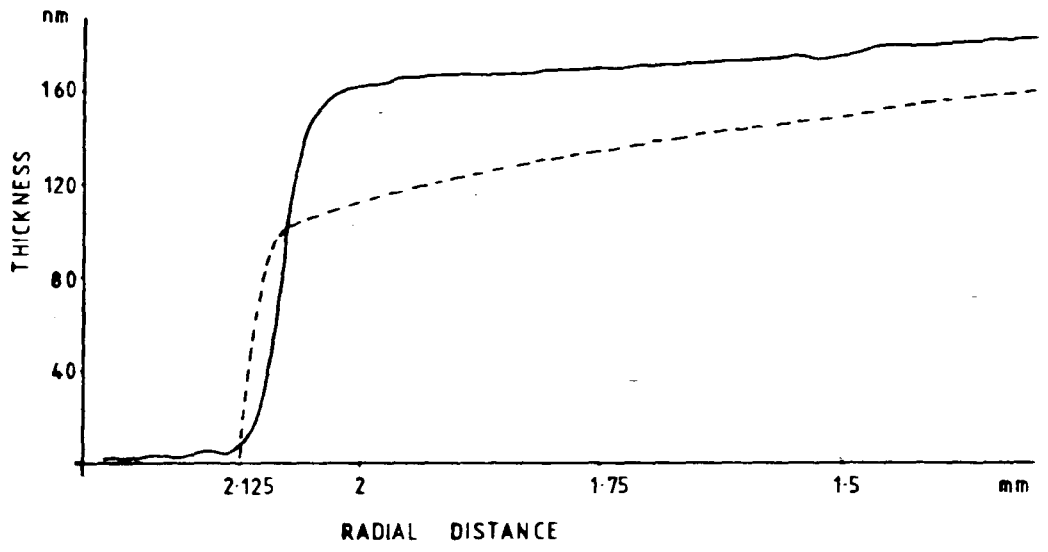


FIGURE [6,11] COMPARISON BETWEEN THEORETICAL AND EXPERIMENTAL LENS PROFILES

F/2 LENS As_2S_3 on $Ti:LiNbO_3$

Attempts to fabricate overlay lenses by reducing the distances between source, profile-mask and substrate to those given in chapter three met with no success. Even very short exposures of the substrate to the source resulted in very thick coatings of the substrate with As_2S_3, so that the desired profiles were never realized. Such a failure may be easily understood by referring to Table [6.1], where a typical film-deposition rate of 3 nm/sec is quoted for a source-to-substrate distance of 200mm. If a simple inverse-square-rule is assumed to apply, the corresponding film deposition-rate at a source-to-substrate distance of 18.4mm (required for fabricating an f/2 lens of 8mm diameter) is 355 nm/sec. Thus, even a ten second exposure would cause 3550 nm to be deposited at the lens centre. An exposure of approximately a half-second would be required to deposit only 185 nm of film, the required thickness at the centre. Even with slower deposition rates the problems were found to be unsurmountable.

Busch et al [55], using the same materials combination and similar fabrication conditions, have reported better results than those reported herein. Better overall profile control was maintained, together with good thickness control. The reasons are two-fold:

1) the crucible used by Busch et al was 18.5mm in diameter compared to the 6.25mm used in the present experiments, and was essentially of the knife-edge type, ie it had virtually no depth. Such a crucible would allow much greater variation of the source-area 'seen' by different points on the substrate, at longer source-to-substrate and source-to-mask distances, thus allowing easier monitoring and better control;

2) the As_2S_3 films deposited by Busch et al were not thermally annealed. As such the refractive index was approximately 2.44, compared to the refractive index of approximately 2.6 obtained in the present experiments. The lowering of the refractive index allows lenses of larger f-number to be fabricated, as observed in chapter three. Larger f-number

lenses imply flatter lens profiles, which again imply longer source-to-substrate distances.

The absence of a thermal annealing stage would likely have led to both physical and optical instabilty of the films, though this was not mentioned by Busch et al. Deliberate photo-annealing was used by them to modify the focal characteristics of the deposited lenses. Such a technique does seem, therefore, to have advantages, although serious questions about the propagation losses, optical stability and mechanical ruggedness of unannealed films remain.

An alternative technique for the manufacture overlay lenses may be to abandon hopes of using the variation of deposited thickness naturally obtained at short source/target distances, and to modify the deposited thickness in other ways. For example a long source-to-substrate distance could be used to obtain a more-or-less uniform deposited film in the absence of a masking aperture. An aperture with a variable diameter, ie an iris, could then be introduced as closely as possible to the substrate. The iris diameter could then be controlled by a computer-driven motor such that true Luneburg profiles could be obtained. An extremely slow, uniform rate of deposition would be probably be required. Detailed calculations on the use of this technique remain to be carried out.

6.5 Conclusions

The lens-profile design criteria established in chapter two and the investigation into vacuum-evaporation in a simple shadow-masking environment presented in chapter three predicted that the accurate formation of inhomogeneous overlay lenses of the Luneburg type would be extremely difficult, for the As_2S_3 on $Ti:LiNbO_3$ combination. Attempts at fabricating such lenses have been reported in this chapter, and the theoretical predictions have been confirmed. It seems impossible to control simultaneously the absolute deposited film thickness and the

profile shape, given the small source-to-substrate and source-to-mask distances required. Although true Luneburg lenses were not obtained, lenses were fabricated and measured. The optical performance of fabricated Luneburg lenses will be reported in chapter eight. An alternative technique for producing Luneburg overlay lenses was proposed. A further problem with the use of As_2S_3 as an overlay material was the very high level of light-scattering and waveguide loss observed.

REFERENCES FOR CHAPTER SIX

[1] Nassau,K and Levinstein,H.J.,
Ferro-electric behaviour of Lithium Niobate, Appl. Phys. Lett., 7(3), 1965, 69-70

[2] Boyd,G.D.,
$LiNbO_3$: an efficient phase-matchable non-linear optical material, Appl. Phys. Lett., 5(11), 1964, 234-236

[3] Midwinter,J.E.,
Assessment of Lithium Metaniobate for non-linear optics, Appl. Phys. Lett., 11(4), 1967, 128-130

[4] Spaight,R.N. and Koerber,G.G.,
Piezoelectric surface waves on $LiNbO_3$, IEEE Trans. Sonics and Ultrason., SU-18(4), 1971, 237-238

[5] Slobodnik,A.J.,
Surface acoustic waves and SAW materials, Proc. IEEE, 64(5), 1967, 581-595

[6] DeLaRue,R.M.,
Integrated acousto-optics: the technology and the competition, from 'New Directions in Guided Wave and Coherent Optics', Ostrowsky,D.B. and Spitz,E. (editors), Martinus Nijhoff, The Hague, 1984

[7] White,J.M., Heydrich,P.F. and Lean,E.G.,
Thin-film acousto-optic interaction in $LiNbO_3$, Electron. Lett., 10(24), 1974, 510-511

[8] Peterson,G.E., Ballman,A.A., Lenzo,P.V. and Bridenbaugh,P.M.,
Electro-optic properties of $LiNbO_3$, Appl. Phys. Lett., 5(3), 1964, 62-64

[9] Glass,A.M., Kaminow,I.P., Ballman,A.A. and Olsen,D.H.,
Absorption loss and photo-refractive index changes in Ti:$LiNbO_3$

crystals and waveguides, Appl. Opt., 19(2), 1980, 276-281

[10] Tsai,C.S.,
Guided-wave acousto-optic Bragg modulators for wide-band integrated optic communications and signal processing, IEEE Trans. Cir. Sys., CAS-26, 1979, 1072-1098

[11] Barakat,R.,
The intensity distribution and total illumination of aberration-free diffraction images, Progress in Optics, I, 1961, 69-105

[12] Van Der Donk,J, Vandewege,J.,
Second-order effects in geodesic lenses, Proc. of the IEE 1st European Conference on Integrated Optics, 1981

[13] Vahey,D.W., Kenan,R.P. and Burns,W.K.,
Effects of anisotropic and curvature losses on the operation of geodesic lenses in Ti:$LiNbO_3$ waveguides, Appl. Opt., 19(2), 1980, 270-275

[14] Schmidt,R.V. and Kaminow,I.P.,
Metal-diffused optical waveguides in $LiNbO_3$, Appl. Phys. Lett., 25(8), 1974, 458-460

[15] Griffiths,G.J. and Esdaile,R.J.,
Analysis of Titanium diffused planar optical waveguides in Lithium Niobate, IEEE J. Quant. Electron., QE-20(2), 1984, 149-159

[16] Noda,J. and Iwasaki,H.,
Impurity diffusion into $LiNbO_3$ and $LiTaO_3$, Proceedings of the 2nd Meeting on Ferroelectric Materials and their Applications, Kyoto,Japan, 24-26 May 1979, 149-154

[17] Kaminow,I.P. and Carruthers,J.R.,
Optical waveguiding layers in $LiNbO_3$ and $LiTaO_3$, Appl. Phys. Lett., 22(7), 1973, 326-328

[18] Minakata,M., Saito,S., Shibata,M. and Miyazawa,S.,
Precise determination of refractive index changes in Ti-diffused $LiNbO_3$ optical waveguides, J.Appl.Phys., 49(9), 1978, 4677-4682

[19] Esdaile,R.J.,
Closed-tube control of outdiffusion during fabrication of optical waveguides in $LiNbO_3$, Appl. Phys. Lett., 33(8), 1978, 733-734

[20] Alferness, R.C., Schmidt,R.V. and Turner,E.M.,
Characteristics of Ti-diffused $LiNbO_3$ optical directional couplers, Appl. Opt, 18, 1979

[21] Tsai,C.S., Kim,B. and El-Akkari,F.R.,
Optical channel waveguide switch and coupler using total internal reflection, IEEE J. Quant. Electron., QE-14(7), 1978, 513-517

[22] Schmidt,R.V. and Alferness,R.C.,
Directional coupler switches, modulators and filters using alternating delta-beta technique, IEEE Trans. Circ. Syst., CAS-26, 1979, 1099-1108

[23] Izutsu,M.,
10 Ghz bandwidth travelling-wave $LiNbO_3$ optical waveguide modulator, IEEE J. Quant. Electron., QE-14, 1978, 394-395

[24] Yi-Yan,A., Andonovic,I., Pun,E.Y.B., and Bjortorp,B.,
Fabrication of periodic Ti:$LiNbO_3$ waveguides by single- and double-diffusion, Appl. Phys. Lett., 43(1), 1983, 19-21

[25] Keil,R. and Auracher,F.,
Coupling of single-mode Ti-diffused $LiNbO_3$ waveguides to single-mode fibers, Opt. Commun., 30, 1979, 23-28

[26] Doughty,G.F., DeLaRue,R.M., Finlayson,N., Singh,J. and Smith,J.F.,
An integrated optical microwave spectrum analyzer (IOSA) using geodesic lenses, Proc. SPIE, vol 369, 1982, 705-710

[27] Bradley,J.C., Hutcheson,L.D., Kellner,A.L., Malarkey,E.C., Mergerian,D. and Pautenius,R.P.,
Geodesic lens performance characteristics, Proc. SPIE, vol 239, 1980, 84-89

[28] Mergerian,D., Malarkey,E.C., Pautenius,R.P., Bradley,J.C., Marx,G.E., Hutcheson,L.D. and Kellner,A.L.,
An integrated optical radio-frequency (rf) spectrum analyzer, Proc. SPIE, vol 239, 1980, 121-127

[29] Kanazawa,M., Atsumi,T., Takami,M. and Ito,T.,
High-resolution spectrum analyzer, Tech. Digest, IOOC '83, Tokyo,1983, Paper 30B3-5, 258-259

[30] Young,P.A.,
Optical properties of vitreous arsenic trisulphide, J. Phys. C: Solid State Phys., 4, 1971, 93-106

[31] Fujiwara,T., Itoh,S. and Okazaki,M.,
Structural model of amorphous As_2S_3, J. Non-cryst. Solids, 45, 1981, 371-378

[32] Watts,R.K., de Wit,M. and Holton,W.C.,
Non-oxide chalcogenide glass films for integrated optics, Appl. Opt., 13(10), 1974, 2329-2332

[33] Stewart,G., Hutchins,R.H. and Laybourn,P.J.R.,
Controlled growth of arsenic trisulphide films for coupling integrated optical devices, J. Phys. D: Appl. Phys., 14, 1981, 323-331

[34] De Neufville,J.P., Moss,S.C. and Ovshinsky,S.R.,
Photo-structural transformations in amorphous As_2Se_3 and As_2S_3 films, J. Non-cryst. Solids, 13, 1973/74, 191-223

[35] Kushibiki,J., Maehara,H., and Chubachi,N.,
Acoustic properties of evaporated chalcogenide glass films, Electron. Lett., 17(9), 1981, 322-323

[36] Macleod,H.A.,
Micro-structure of optical thin films, Proc. SPIE, 325, Optical Thin Films, 1982, 21-28

[37] Zallen,R.,
'Physics of Amorphous Solids', Wiley and Sons, New York 1983

[38] Suhara,T., Nisihara,H., and Koyama,J.,
Design of high efficiency waveguide holograms, IOOC '77, paper B4-3, 1977, 235-238

[39] Andriesh,A.M., Bykovskii.Yu, Kolomeiko,E.P., Makovkin,A.V., Smirnov,V.L. and Schmal'ko,A.V.,
Waveguide structures and functional elements of integrated optics systems based on volume holographic gratings in thin As_2S_3 films, Sov. J. Quant. Electron., 7(3), 1977, 347-352

[40] Noda,J., Zembutsu,S., Fukunishi,S. and Uchida,N.,
Strip-loaded waveguide formed in a graded-index $LiNbO_3$ planar waveguide, Appl. Opt., 17(12), 1978, 1953-1958

[41] Korsakov,V.V., Fateev,V.A. and Tsukerman,V.G.,
Electro-optic modulation of laser beams in thin-film optical waveguides of amorphous chalcogenides, Sov. Tech. Phys. Lett., 7(2), 1981, 69-70

[42] Ohmachi,Y.,
Acousto-optical light diffraction on thin-films, J. Appl. Phys., 44(9), 1973, 3928-3933

[43] Uchida,N. and Niizeki,N.,
Acousto-optic deflection materials and techniques, Proc. IEEE, 61(8), 1973, 1073-1092

[44] Loh,K.W., Chang,W.S.C. and Becker,R.A.,
Convolution using guided acoust-optical interaction in As_2S_3

waveguides, Appl. Phys. Lett., 28(3), 1976, 109-111

[45] Yaji,T., Kurita,S., Tejima,Y., Horinouchi,A. and Koyanagi,M.,
Integrated optical devices using amorphous As_2S_3 thin-film, IEEE Trans. CHMT, CHMT-6(1), 1983, 134-135

[46] Suhara,T., Kobayashi,K., Nisihara,H. and Koyama,J.,
Graded-index Fresnel lenses for integrated optics, Appl. Opt., 21(11), 1982, 1966-1971

[47] Handa,Y., Suhara,T., Nisihara,H. and Koyama,J.,
Micro-gratings for high efficiency guided beam deflection fabricated by electron-beam direct writing techniques, Appl. Opt., 19(16), 1980, 2842-2847

[48] Smith,W.R., Gerard,H.M., Collins,J.H., Reeder,T.M. and Shaw,H.J.,
Design of surface wave delay lines with interdigital transducers, IEEE Trans. Micro. Theory Tech., MTT-17(11), 1969, 865-873

[49] Doughty,G.F.,
'Aspheric geodesic lenses for an integrated optic spectrum analyzer', Ph.D. Thesis, University of Glasgow, 1983

[50] Holman,R.L., Cressman,P.J. and Anderson,J.A.,
Processing and properties of diffused Lithium Niobate planar optical waveguides, Ferroelectrics, 27, 1980, 77-80

[51] Stulz,L.W.,
Titanium in-diffused $LiNbO_3$ optical waveguide fabrication, Appl. Opt., 18(12), 1979, 2041-2044

[52] Vahey,D.W.,
In-plane scattering in $LiNbO_3$ waveguides, Proc. SPIE, 176, Guided-Wave Optical Systems and Devices II, 1979, 62-69

[53] Barr and Stroud Technical Literature, Barr and Stroud Limited, Anniesland, Glasgow, G13

[54] Tien,P.K.,
Light waves in thin-films and integrated optics, Appl. Opt., 10(11), 2395-2413
(1971)

[55] Busch,J.R., Wood,V.E., Kenan,R.P. and Verber,C.M.,
Evaporated As_2S_3 Luneburg lenses for $LiNbO_3$:Ti optical waveguides, NASA Conference Publication 2207, NASA Scientific and Technical Information Branch, 1981, 251-261

[56] Canali,C., Carnera,A., Mazzoldi,P. and De La Rue,R.M.,
$LiNbO_3$ optical waveguide fabrication by Ti indiffusion and proton-exchange: process, performance and stability, Proc. SPIE, 517, Integrated Optical Circuit Engineering, 1984, 119-127

[57] Silva, W.J. and Bulmer, C.H.
Compositional effects on Lithium Niobate substrates
Proc. SPIE, 578, Integrated Optical Circuit Engineering II, 1985, 19-21

CHAPTER SEVEN

WAVEGUIDE UNIFORMITY

GEODESIC LENS WAVEGUIDES

7.1 Introduction

The geodesic lenses used in the present study were fabricated by G.F. Doughty et al to function as collimation and Fourier-transforming elements in an integrated optical spectrum analyser (IOSA) [1] . Doughty [2] reported on many aspects of design and fabrication of integrated optical geodesic lenses. Singh [3] investigated the scattering properties of the Ti-diffused waveguides which were to guide light in both the planar and lens regions of the IOSA. However Singh's work was largely confined to planar waveguides. This chapter will briefly discuss the problem of obtaining uniformity of Ti-diffused waveguides situated in the lenses themselves. Theoretical calculations performed by the author indicate that the problem of achieving waveguide uniformity, which has been largely ignored in the literature, is a significant one that is likely to cause a degradation in lens performance. Alternative waveguide technologies will be proposed which should improve uniformity.

7.2 Factors affecting transmission through lenses

Vahey et al [4] have pointed out that aberrations are not the only phenomena that can affect lens performance in integrated optics applications. Aberrations are essentially phase-front distortions. However the amplitude profile of a wave-field propagating through a lens can also be distorted in ways that may or may not be favourable for the optical properties of the lens. It was observed in chapter five that a beam with a truncated Gaussian amplitude profile propagates much more smoothly in a lens than a single truncated plane-wave, resulting in improved focal plane characteristics, particularly in terms of sidelobe levels. A beneficial shaping of the beam may equivalently be

carried out by controlling the transmission function of the lens itself, a process known as apodisation. In integrated optical geodesic lenses fabricated on anisotropic substrates two mechanisms can be identified which affect the lens transmission properties [4] :

(i) Anisotropy, in the form of birefringence, can give rise to leaky modes which cause power loss. Birefringence in a curved anisotropic substrate is a function of position, as well as direction. Leaky modes arise as a result of coupling between TE and TM polarizations in non-axial propagation directions. In y-cut $LiNbO_3$ high propagation loss can occur for the TE polarization, as shown by Sheem and co-workers [5].

(ii) The curvature of the substrate leads to radiation loss associated with the bending of the waveguide across the lens. In aspheric lenses the curvature is a position-dependent function.

Using a simple model, Vahey et al found that leaky-mode losses are dominant in cases of good field confinement, far from cut-off. The losses are reasonably small in such cases, of the order of 3db/cm to 7db/cm. Leaky-mode loss becomes small for propagating modes approaching cut-off. Lenses with low f-numbers experience greater leaky-mode losses than those with high f-numbers, as would be expected since significant portions of the wave-field then propagate in directions far from the axis. For the same reason, leaky-mode losses are lowest for light propagating through the centre of the lens and highest at the margins. The net spatial distribution of the losses is such that their effect is beneficial to lens operation, with the focal side-lobe intensity being reduced.

On the other hand, losses associated with curvature were large and dominant in cases of poor field confinement, with the fundamental mode close to cut-off. With good field confinement, far from cut-off, curvature losses become very small. Losses due to curvature are greatest for light travelling through the centre

of the lens and smallest for light traversing the marginal parts of the lens. If curvature losses are large, the performance of the lens may be severely degraded as the losses cause an increase in sidelobe intensity.

The transmission functions of geodesic lenses manufactured in $LiNbO_3$ substrates are, therefore, non-uniform and complicated owing to the influence of curvature and anisotropy, both of which are non-uniform, position-dependent functions.

The curvature of the substrate also gives rise to non-uniformity in Ti-diffused waveguides. Non-uniformity may then be a third mechanism affecting the amplitude transmission function of geodesic lenses. Furthermore, non-uniformity may also give rise to phase-distortions. The degree of non-uniformity of titanium films on geodesic lens substrates is investigated in the following section.

7.3 Profiles of films deposited on geodesic lenses

It was shown in chapter three that a directed-surface source was a good approximation to the type of source used in vacuum evaporation chambers. As the name implies, such a source has directional emission properties. The thickness of film material deposited on a target surface is then a function of the source emission properties and the spatial orientation of the surface. It would be expected that thin-films deposited on geodesic lenses having strongly curved profiles would exhibit a considerable variation of thickness over the profiles.

A simple geodesic lens to analyse in terms of deposited film uniformity is a lens possessing a spherical inner geometry and a toroidal rounded-edge to bridge between the inner region and the ambient, planar waveguide. Such a lens is easily described analytically, and has previously been described in this thesis. A diagram of the generating curve of such a lens is shown in Figure [7.1] . The inner region has a radius of curvature R_s and the toroidal region has a radius of curvature a.

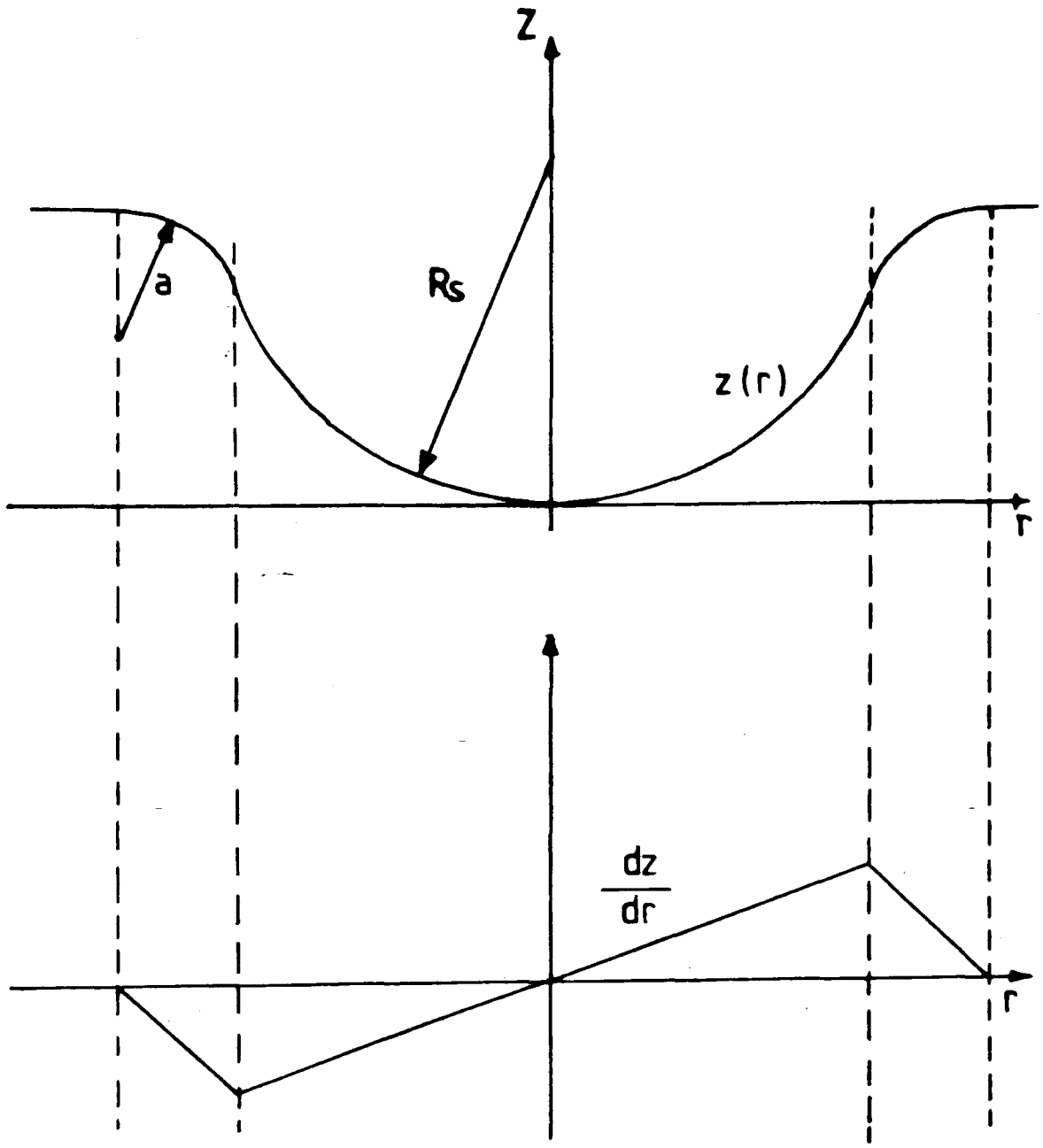


FIGURE [7·1] Behaviour of z(r) and dz/dr for a spherical geodesic lens with toroidal edge-rounding

The dimensions and position of the toroidal region are chosen so that the tangents at the intersection points with both the planar substrate and the inner lens are matched. The profile may then be described as a smooth function of the axial coordinate r. The first derivative of the generating curve is also shown in Figure [7.1]. It is evident that the first derivative function is not smooth. The four intersection points on the profile are cusps, separating regions having distinctly different gradient.

In many applications the substrate is required to support two lenses, one for collimation and another for Fourier transforming purposes. Vacuum evaporation units used for deposition of thin films onto such substrates are often unsophisticated, and the positon of substrate and source often has to be held constant throughout deposition. The procedure adopted to obtain a degree of uniformity in such a case is to make the source-to-substrate distance as large as possible, and to position the substrate so that the lenses are symmetrically aligned with respect to the source. Just how uniform a film is deposited may be investigated using the concept of the directed-surface source.

The geometry of the system is shown in Figure [7.2]. A directed-surface source emits an amount of evaporated material in a given direction which depends on the variable θ. The source is parallel with the horizontal axis and with the planar portion of the substrate. The deepest point of the lens lies on the horizontal axis and the profile function, including the planar portion, may be described by a function z(r). The thickness variation along one axis only is therefore being considered. The planar portion of the substrate is situated a distance z_p from the axis, and the source is situated a distance z_o from the origin.

The directed-surface source causes a spatial distribution of deposited film thickness on the surface z(r). At a typical point P_v such as shown in Figure [7.2] the thickness is given by

on the axial section

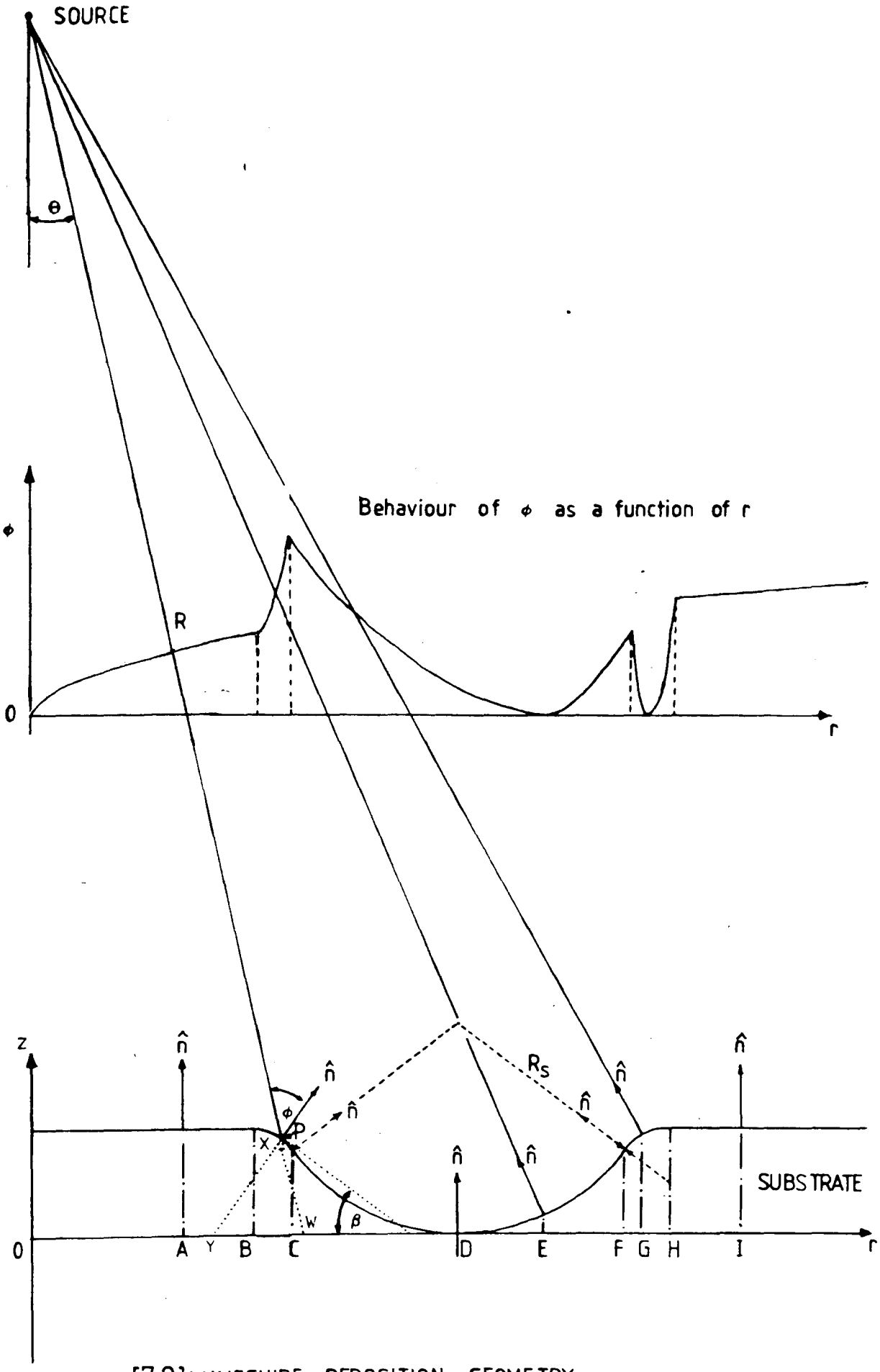


FIGURE [7,2] WAVEGUIDE DEPOSITION GEOMETRY

$$t(r) = \frac{k \cos\Theta(r) \cos\phi(r)}{R^2(r)} \qquad (7.1)$$

where a line drawn between the source and P defines the direction of emission, Θ is the angle between the source normal and the direction of emission, ϕ is the angle between the surface normal and the direction of emission and R is the distance between source and substrate along the direction of emission. Equation (7.1) was previously encountered in chapter three and is known as Knudsen's cosine law of emission. k is a constant factor under conditions of source equilibrium. At the point on the substrate directly below the source, the deposited thickness obtains its maximum value and is given by

$$t_o = \frac{k}{R_o^2} \qquad (7.2)$$

The deposited thickness at the other points on the surface may be normalised to the maximum value:

$$t' = t/t_o = \frac{\cos\Theta \cos\phi R_o^2}{R^2} \qquad (7.3)$$

The variables R and Θ are defined in terms of the known quantities $z(r)$, r , and z_o:

$$\cos\Theta = \cos\left(\tan^{-1}\left(\frac{r}{z_o - z(r)}\right)\right) \qquad (7.4)$$

and

$$R^2 = (z_o - z(r))^2 + r^2 \qquad (7.5)$$

The remaining variable ϕ may be determined by the same quantities and the first derivative of the profile. In Figure [7.2] the tangent to the lens at point P is shown intersecting the horizontal axis at an angle β . The magnitude of this angle is given by :

$$|\beta| = \tan^{-1}\left(\left|\frac{dz}{dr}\right|\right) \qquad (7.6)$$

The angles that make up triangle WXY define ϕ:

$$\phi = \Theta + \beta \qquad (7.7)$$

The relationship (7.7) will hold for all points with negative slope, ie for points between the origin and the centre of the lens. For points with positive slope the magnitude of ϕ is given by

$$|\phi| = |\Theta - \beta| \qquad (7.8)$$

Substituting for Θ according to (7.4) and for β according to (7.6), the expression for ϕ becomes

$$\phi = \tan^{-1}\left(\frac{r}{z_o - z(r)}\right) \pm \tan^{-1}\left(\frac{dz}{dr}\right) \qquad (7.9)$$

Formulae (7.1) to (7.9) apply to any geodesic lens substrate, whether of spherical or aspheric geometry. At large source-to-substrate distances , R and Θ vary slowly as a function of r. The variable ϕ then exerts the strongest influence on the deposited thickness. The behaviour of ϕ for a spherical geometry lens may be determined by considering the nine points A-I in Figure [7.2]. The surface normals are shown for convenience at each of the points.

The behaviour of ϕ is sketched in Figure [7.2]. At points A and I corresponding to typical points on the plane surface of the substrate, and at point D at the centre of the lens, the first derivative of the function is zero, and $\phi = \Theta$ as a result. ϕ behaves as the inverse tangent function with respect to r in these areas. Point B is the first point of intersection between the plane and the toroidal rounded-edge. Due to the behaviour of the first derivative at this point, as seen in Figure [7.1], a

cusp occurs in the function ϕ .As the point of interest P moves from B to the point of intersection C between the toroidal rounded-edge and the inner portion of the lens, ϕ increases, reaching a maximum at C. A second cusp occurs at C and o begins to decrease thereafter. ϕ is non-zero at the centre of the lens but eventually becomes zero at point E where the surface normal and the direction of emission coincide. ϕ begins to rise again beyond E until a second maximum occurs at point F, where the inner portion of the lens and the toroidal rounded-edge again intersect. This maximum is again a cusp, following the behaviour of the first derivative of the profile function. Beyond point F ϕ decreases, reaching zero again at point G. ϕ increases thereafter until the final cusp at H occurs. The function ϕ then resumes inverse tangent behaviour.

The variation of ϕ as described above is reflected in the variation of deposited film thickness on geodesic lenses. Figures [7.3] and [7.4] show the thickness variation across the spherical lens with toroidal edge-rounding shown in chapter two. The lens diameter is 10.462 mm and the centre of the lens is situated at an offset distance from the point directly below the source, a situation likely to occur when two lenses are present on one waveguide. In Figure [7.3] the offset distance is 10.0 mm. Three curves are plotted in Figure [7.3], calculated for source-to-substrate distances of 49.12 mm, 99.12 mm and 199.12 mm. The behaviour of each curve is closely related to the behaviour of the curve previously shown in Figure [7.2]. The most extreme behaviour is, as expected, manifested by the curve calculated for the shortest source-to-substrate distance and may be considered in detail. At the edge of the lens, closest to the point directly below the source, the film thickness falls to 98% of its maximum value. The thickness then rapidly falls to under 90% of its maximum value over the toroidal edge-rounding region and then settles again, changing very slowly over the major portion of the lens. A small increase in thickness occurs as the toroidal region is encountered a second time, until the substrate normal points directly towards the source. Beyond this point, the thickness rapidly falls off, down to 84% of its maximum value until the

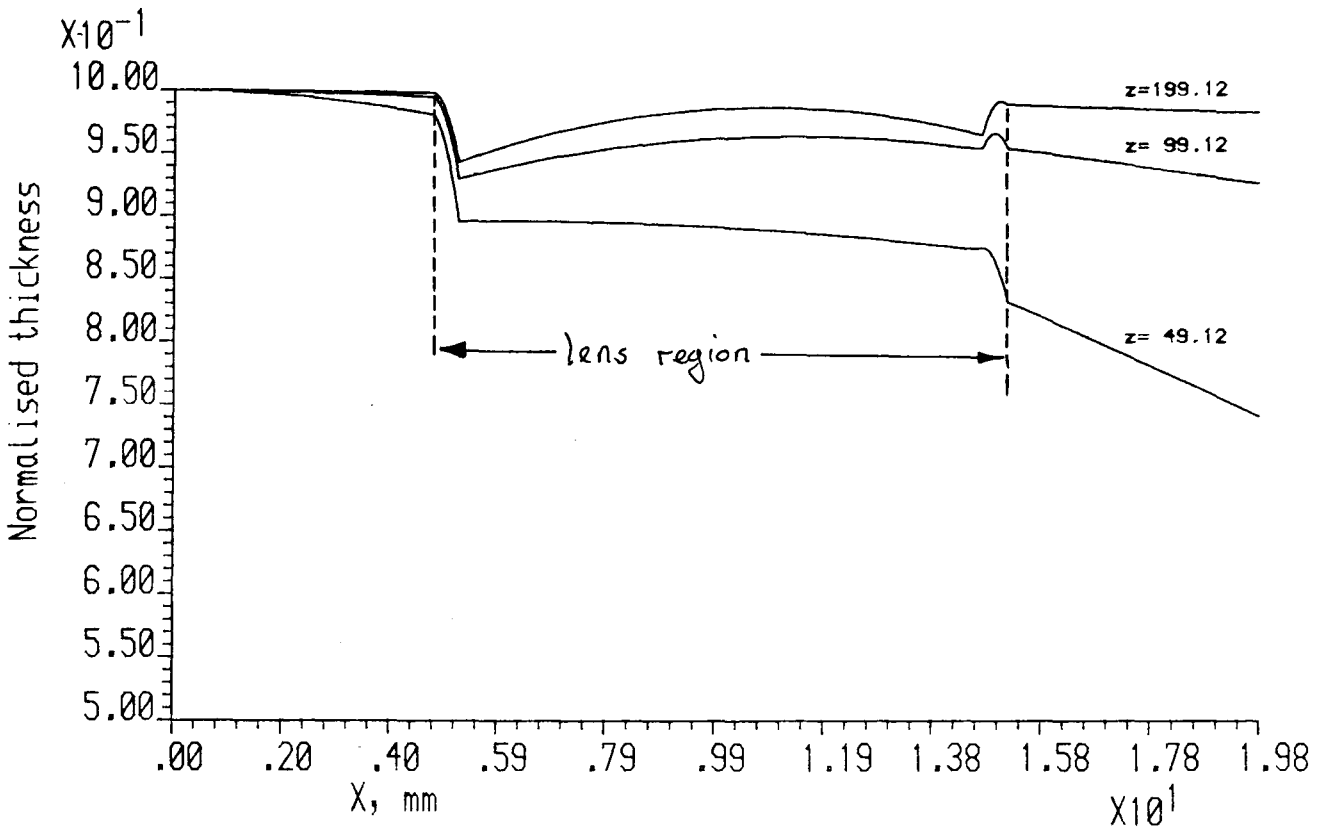


FIGURE [7,3]

Film deposition thickness on LiNbO3 substrate

incorporating spherical geodesic lens

Lens parameters:

Focal length= 52.3mm

Full aperture= 10.462mm; useful aperture= 4.0mm

Deposition parameters:

source to substrate distances= z mm

horiz. distance from source to lens centre= 10.00 mm

ASSUMPTIONS: directed surface source

sticking coefficient of unity

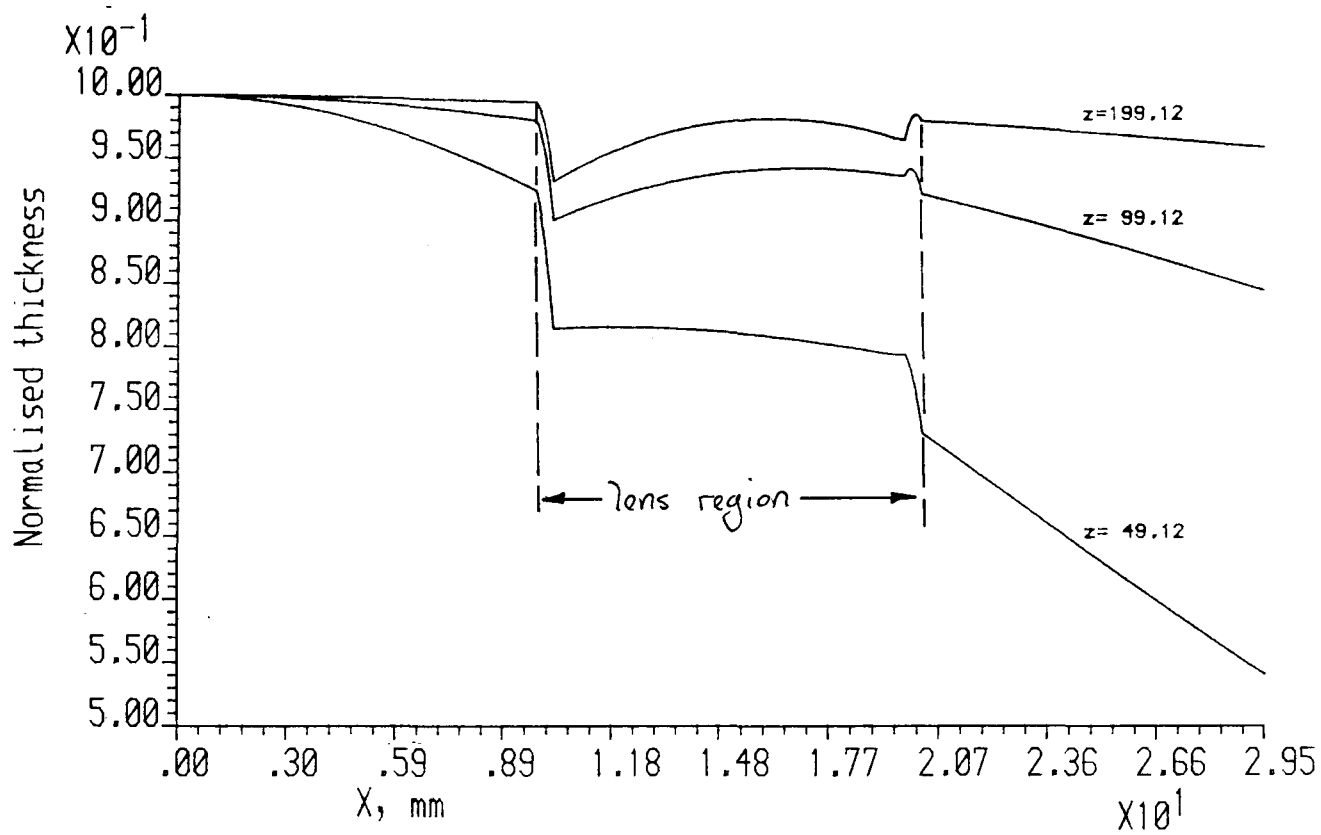


FIGURE [7,4]

Film deposition thickness on LiNbO3 substrate

incorporating spherical geodesic lens

Lens parameters:

Focal length= 52.3mm

Full aperture= 10.462mm; useful aperture= 4.0mm

Deposition parameters:

source to substrate distances= z mm

horiz. distance from source to lens centre= 15.00 mm

ASSUMPTIONS: directed surface source

sticking coefficient of unity

ambient planar substrate is encountered, at which point the rate of change of thickness as a function of position across the substrate decreases.

For source-to-substrate distances of 99.12 mm and 199.12 mm the behaviour is similar to the 49.12 mm case, except that the deposited thickness actually increases as a function of distance from the point directly under the source, over large portions of the inner section of the lens. A greater degree of uniformity is obtained for larger source-to-substrate distances, as would be expected, but a maximum variation of 6% is still obtained for the 199.12 mm case. The rate of change of thickness in local areas is also quite substantial, even for large source-to-substrate distances.

The calculations were repeated for a larger lens offset distance of 15.0 mm. The results are shown in Figure [7.4]. Qualitatively similar behaviour to that of Figure [7.3] is obtained , though the actual variation in deposited film thickness over the substrate is larger. Again, more-or-less 'flat' thickness profiles are obtained over the inner portion of the spherical lens with toroidal edge-rounding.

'Flat' behaviour is certainly not obtained, however, over aspheric geodesic lenses such as were specified in chapter two. The curvature variation of the aspheric lenses is much stronger than that of lenses based on spherical geometry, and this is reflected in the variation of deposited film thickness. Figures [7.5] and [7.6] show deposited thickness curves calculated for the aspheric geodesic lens investigated by Doughty [2] and the present author under similar conditions to those used for Figures [7.3] and [7.4]. The deposited film profiles are considerably different from those of the spherical lens with edge-rounding. At a source-to-substrate distance of 48.69 mm and for a centre-offset distance of 10.0 mm, the deposited thickness falls to below 75% of the maximum value at the first boundary between the edge-rounding region and the inner lens. For the same source-to-substrate distance, and an offset distance of 15.0 mm the minimum

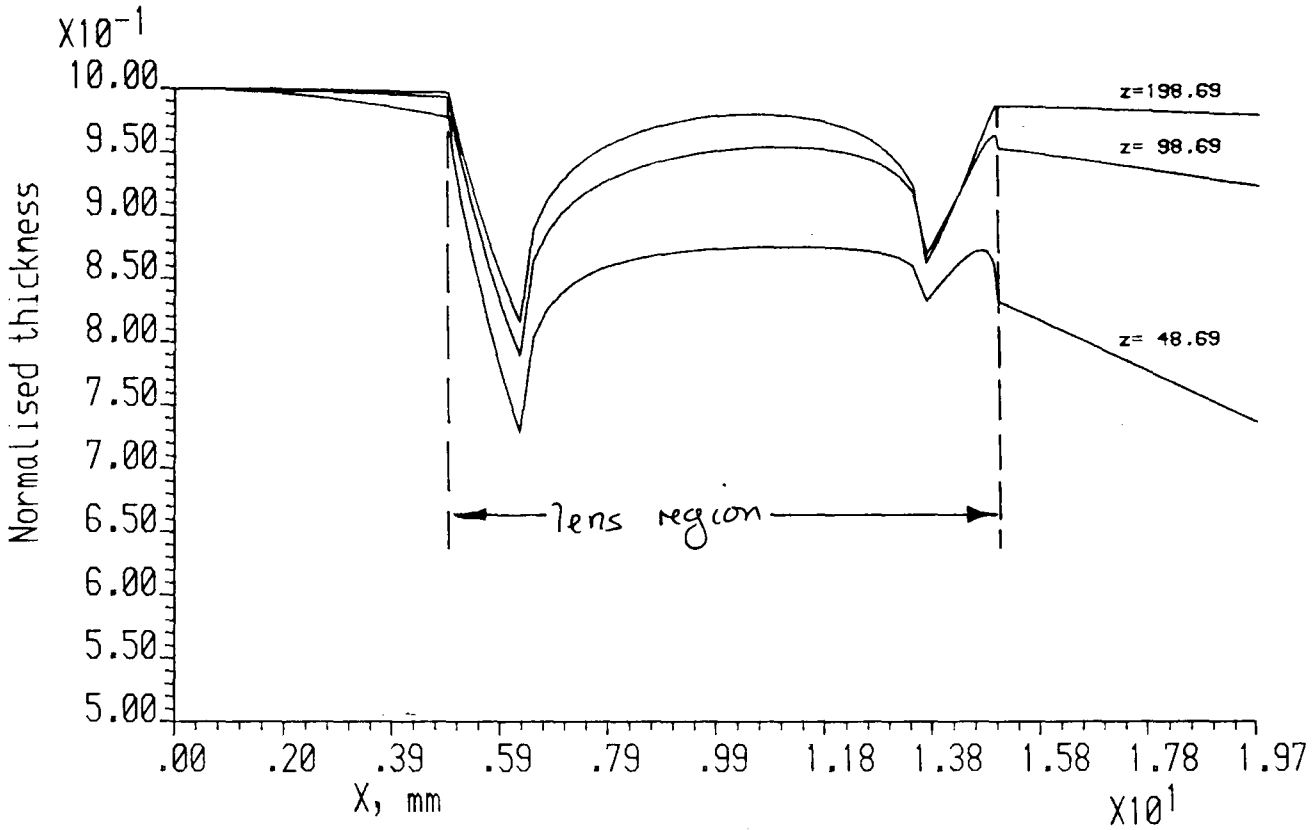


FIGURE [7,5]

Film deposition thickness on LiNbO3 substrate

incorporating Sottini geodesic lens

Lens parameters:

Focal length= 18.5mm

Full aperture= 10.0mm; useful aperture= 7.4mm

Deposition parameters:

source to substrate distances= z mm

horiz. distance from source to lens centre= 10.00 mm

ASSUMPTIONS: directed surface source

sticking coefficient of unity

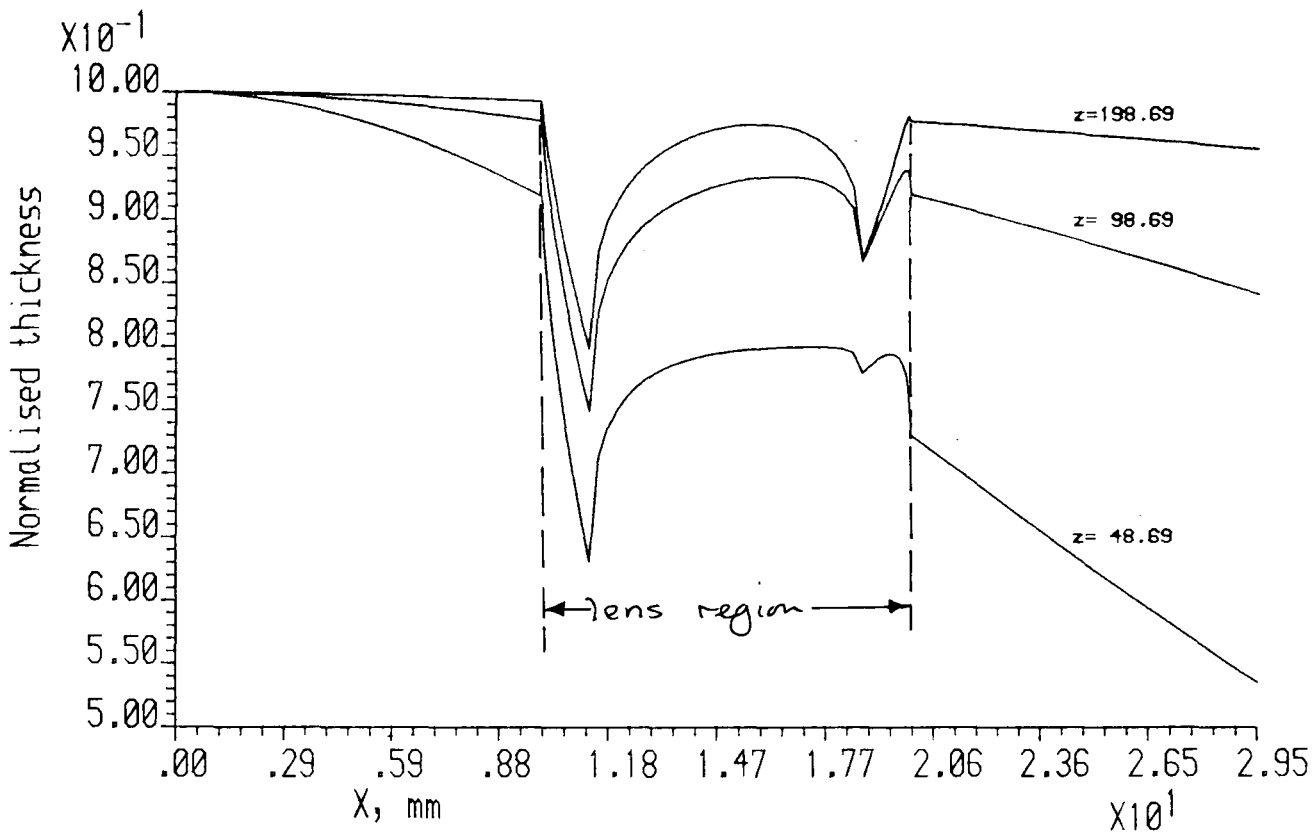


FIGURE [7,6]

Film deposition thickness on LiNbO3 substrate

incorporating Sottini geodesic lens

Lens parameters:

Focal length= 18.5mm

Full aperture= 10.0mm; useful aperture= 7.4mm

Deposition parameters:

source to substrate distances= z mm

horiz. distance from source to lens centre= 15.00 mm

ASSUMPTIONS: directed surface source

sticking coefficient of unity

deposited thickness falls below 65%. Even for a long source-to-substrate distance of 198.69 mm, the minimum thickness is only 82% for an offset distance of 10.0 mm, and less than 80% for an offest distance of 15.0 mm. Furthermore, the rate of change of deposited thickness across the lens is complicated.

The consequences of such deposited film profile behaviour for the optical performance of these lenses are difficult to ascertain precisely. Only the profile variation along a meridional curve has been calculated, though the variation at other points could also be calculated relatively easily. It would then be possible, presumably, to transform the variation of thickness over the lens into an equivalent effective refractive index perturbation which could be incorporated in ray-tracing or BPM calculations. Without going into such detail, however, it is still possible to make some general observations.

1) The optical path length of rays traversing the lens would be perturbed by the effective index variation arising from the deposited thickness variation. The focal characteristics would be affected by these perturbations of phase. To what extent they would be affected remains to be determined, though one might expect the effects to be small since the refractive index change induced by the titanium diffusion process is small anyway.

2) If the cusps in the thickness profile were to manifest themselves in the refractive index profile after diffusion (which is, perhaps, unlikely) discontinuities would occur in the first derivative of the refractive index function. In these areas strong and unpredictable refraction would probably occur.

3) The film thickness variations could result in the waveguide falling below cut-off in certain areas, if the waveguide at the point of maximum thickness is not very far into the single-mode regime.

4) The thickness variation could lead to losses within the lens.

If all the above observations are valid, waveguide non-uniformity would be expected to introduce both amplitude- and phase-distortions into the propagating optical field.

Experimental confirmation of the theoretical results would be desirable. Improved uniformity of deposition could possibly be obtained using well-established thin-film techniques such as rotation of one, or both, of the substrate and the source; the use of two or more simultaneous sources; or the use of sources shaped such that the source curvature would complement the curvature of the lenses. All such solutions would require quite complicated mathematical analyses, and simpler solutions based on alternative waveguide technologies that do not involve evaporation or sputtering might be desirable. One such technology would be the LiO_2 high-temperature out-diffusion technique mentioned in chapter six. This technique, unfortunately, creates very deep optical waveguides which are not compatible with surface acoustic waves such as are used in the IOSA. A better solution might be the recently-developed technique of proton-exchange. Very recent results indicate that proton-exchange using dilute melts may be capable of superior performance to titanium-diffusion as a method of fabricating waveguides, with 0.5-1.5dB/cm losses, -40dB levels of in-plane scatter at 0.5^{o} and high resistance to optical damage being reported [6]. The refractive index change obtained with the proton-exchange process is high, however, and this would complicate matters in geodesic lenses, since the focusing mechanism could no longer be considered to arise purely from the geometrical, rather than optical profile.

7.4 Conclusions

It has been suggested in this chapter that the problem of obtaining waveguide uniformity be added to the problems of anisotropy and curvature-loss in the study of factors affecting the optical performance of geodesic lenses.

Whilst occassional comments on this subject have been made in the literature of geodesic lenses, no work has been published of either a theoretical or experimental nature. One suspects that the assumption that reasonable uniformity is obtained at large source-to-substrate distances is a widely-held one, although the present work shows that assumption to be manifestly false. The lack of uniformity in the deposited thickness profile which would be expected to appear in the diffused profile also (to a lesser extent) would be undesirable in itself; but the rapid variation of the thickness in certain areas of the profile could considerably worsen the optical performance of these lenses.

REFERENCES FOR CHAPTER SEVEN

[1] Doughty,G.F., DeLaRue,R.M., Finlayson,N., Singh,J. and Smith,J.F.,
Integrated optical microwave spectrum analyser (IOSA) using geodesic lenses, Proc. SPIE, 369 Max Born Conference, Edinburgh, Scotland, 1982, 705-710

[2] Doughty,G.F.,
'Aspheric Geodesic Lenses for an Integrated Optical Spectrum Analyser', Ph.D. Thesis, University of Glasgow, 1983

[3] Singh,J.,
'Studies Concerning the Dynamic Range of the Integrated Optic Spectrum Analyzer', Ph.D. Thesis, University of Glasgow, 1983

[4] Vahey,D.W., Kenan,R.P. and Burns,W.K.,
Effects of anisotropic and curvature losses in Ti:$LiNbO_3$ waveguides, Appl. Opt., 19(2), 1980, 270-275

[5] Sheem,S.K., Burns,W.K. and Milton,A.F.,
Leaky-mode propagation in Ti-diffused $LiNbO_3$ and $LiTaO_3$ waveguides, Optics Lett., 3(3), 1978, 76-78

[6] Wong,K.K., Parsons,N.J., Oldroyd,A.R. and O'Donnell,A.C.,
High-quality optical waveguides in $LiNbO_3$ by dilute-melt proton-exchange, to be presented at IOOC/ECOC '85, Venice, October 1985

CHAPTER EIGHT

OPTICAL EXPERIMENTS

CHAPTER EIGHT

OPTICAL EXPERIMENTS

8.1 Experimental objectives

Having investigated various problems concerning the design, analysis and fabrication of inhomogeneous integrated optical lenses, the ultimate test is to determine how the fabricated lenses perform under a variety of experimental conditions. No single number characterizes the 'quality' of a lens. Several tests are required to establish the suitability of a lens for the tasks it will be required to perform. Integrated optical lenses are usually required to perform well in the back-focal-plane as opposed to providing good imaging at a given magnification. Attention may therefore be confined to the Gaussian image region of a single point. In terms of geometric aberrations, defocus and spherical aberrations are the quantities of interest. From the point of view of physical optics, (incorporating wave aberration and diffraction phenomena) image spot dimensions, sidelobe levels, sidelobe decay, image symmetry and scattering levels are important quantities.

Four lens systems were available for investigation, viz.

(i) a spherical geodesic lens having a toroidal edge-rounding region bridging the inner portion of the lens and the ambient waveguide;

(ii) an overlay Luneburg lens;

(iii) a single aspheric geodesic lens of equivalent Luneburg lens design;

and (iv) a two-lens aspheric geodesic system expected to form the basis for an integrated optical radio-frequency spectrum

analyzer [19].

The objectives of the experiments were to master and, if possible, to improve upon techniques for optical characterisation of these lenses, to use these techniques to characterise each type of lens, and thereby to form a judgement, based on optical criteria alone, on the relative merits of each system.

8.2 Lenses selected for study

The physical characteristics of the four lens systems will now be dealt with in turn. In order to evaluate the performance of fabricated lenses representative of the several types dealt with in this thesis, a broad initial characterisation was carried out to establish those lenses that were likely to yield best results. If, for example, excessive in- or out- of plane scattering levels were detected in the lenses or in the surrounding waveguides, detailed characterisation of the optical performance was not necessary, as the focal pattern was always degraded accordingly.

Four structures supporting a total of five lenses (one substrate carried two lenses) were chosen for detailed study. These will now be described and labelled.

(i) LENS SL1

Lens SL1 was a spherical geodesic lens having a toroidal edge-rounding region. It was fabricated by G.F. Doughty and co-workers of the Department of Electronics and Electrical Engineering at the University of Glasgow for the Microwaves and Acoustics Group at the Marconi Research Centre, Chelmsford, England who were the industrial sponsors of the present author. The author, assisted by Mr A. Hodkin of Marconi Research, created a Ti-diffused waveguide on the lens and supporting substrate and characterised the optical performance of the resulting structure. The structure was intended to operate as the Fourier transforming element in an acousto-optic spectrum

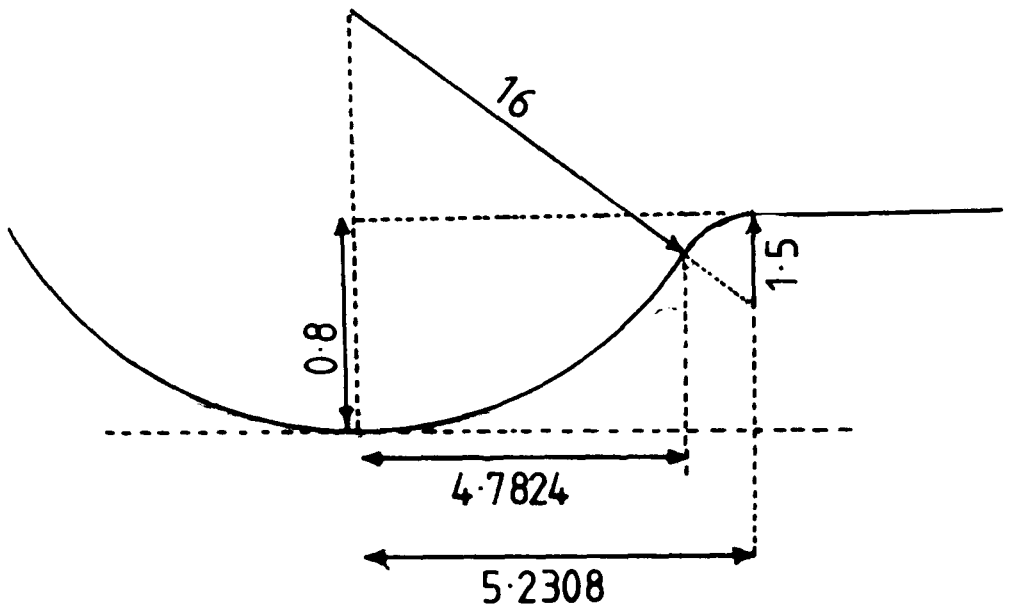
analyzer of hybrid design (the light source, collimation element and detectors were situated off the substrate, with only the acousto-optic interaction region and the transforming lens being 'integrated'). For the purposes of investigating the lens resolving power in conjunction with the acousto-optic interaction, a bank of rf/SAW transducers were incorporated onto the structure after the optical waveguide had been formed.

The physical dimensions of the lens and the supporting substrate are shown in Figure [8.1]. The central spherical depression had a profile radius of curvature of 16.0 mm, with the projection of the depression extending to a radius of 4.782 mm in the waveguide plane. The edge-rounding region was of toroidal design with a radius of curvature of 1.5 mm, and extended a further radial distance of 0.448 mm in the waveguide plane, thus giving the lens an overall radius, in the waveguide plane, of 5.231mm. The paraxial focal length of the lens was intended to be 50.0mm. Precise ray-tracing using Southwell's method indicated that the paraxial focal-length would in fact be 50.9mm.

It is difficult to obtain high-quality boules of $LiNbO_3$ of diameter greater than approximately 50.0mm. It was accepted, therefore, that the converging beam in the image region would have to propagate partly in air. As can be seen from the Figure, the maximum distance from the centre of the lens to the edge of the substrate was only 27mm. Propagation outside the substrate results in a reduced paraxial focal length in accordance with the bending of the rays at the $LiNbO_3$/air interface. Using Snell's Law, and considering only paraxial quantities, the modified focal length of this lens becomes:

$$f' = (f-27)/n_{eff} + 27mm \qquad (8.1)$$

where n_{eff} is the effective refractive index of the light wave in the waveguide and f and f' are measured in mm. At a wavelength of 0.6328um the effective refractive index of the fundamental TE mode of the ambient waveguide was measured to be 2.2065. Using f = 50.9mm the modified focal length then turns out to be 37.8mm.



(a)

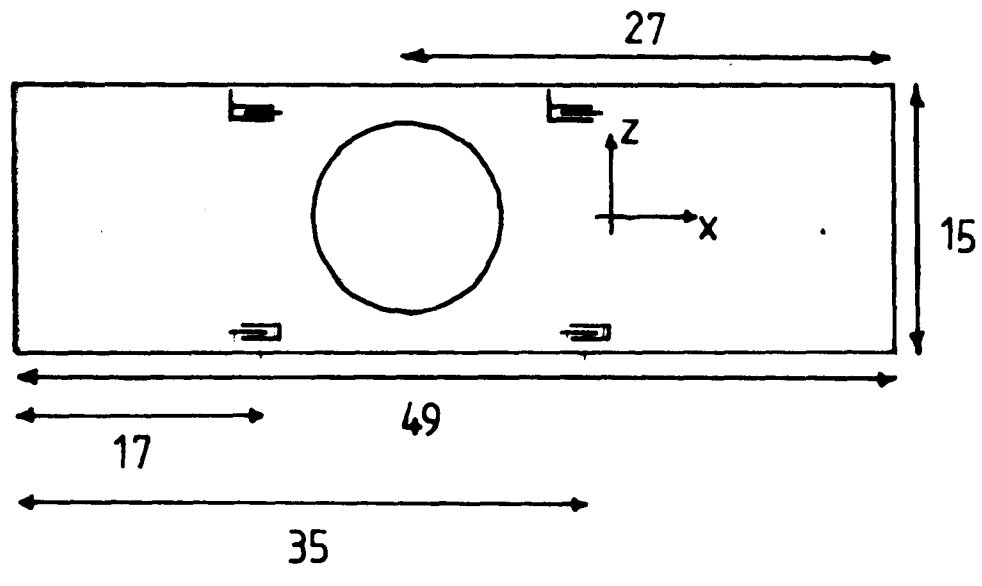
FIG [8,1]

(a) lens dimensions

(b) substrate dimensions, showing SAW transducer bank positions

Notes

1. All dimensions in mm
2. Substrate material:
   y-cut LiNbO
3. Waveguide formed by
   Ti-diffusion



These are physical, not optical distances. A benefit of focusing outside the substrate was that scattering was not too destructive in the far-field. The scattering arises mainly at the waveguide-air interface and is associated with the mean surface-roughness of this interface. Focusing outside the waveguide caused the scattered light to diverge rapidly, while the unscattered light was still caused to converge by the action of the lens.

(ii) LENS LL1

Lens LL1 was an overlay lens approximation to a true Luneburg lens, fabricated with As_2S_3 as the lens material. It was fabricated on a y-cut $LiNbO_3$ substrate on which a Ti-diffused waveguide had previously been formed. The diameter of the lens was 4.25 mm. The thickness of As_2S_3 film deposited at the centre was approximately 185 nm, which, if the lens were a true Luneburg lens, would yield a full aperture f-number of f/2, from the calculations carried out in chapter two. The focal length of the true Luneburg lens would be 8.5 mm in such a case. The profile of the lens was shown in chapter six, and it was observed there that the lens profile over most of the full aperture was thicker than the true Luneburg lens. The lens would therefore perform in a similar fashion to the approximate profile studied, using ray-tracing methods, in chapter four. Considerable spherical aberration would therefore be expected, together with a negative defocus, ie the focus would be expected to be closer to the lens than in the true Luneburg lens.

(iii) LENS GL1

GL1 was an aspheric geodesic lens designed using the method of Sottini, reported in chapter two and fabricated by G.F. Doughty and co-workers. The author's task was to characterize the optical performance of the lens. The central depression (the focusing region) had a diameter of 7.4mm and the edge-rounding region brought the total lens diameter up to 10.0mm. The design focal length of this lens was 18.5mm, implying an effective f-

number of f/2.5 with the central focusing region fully utilised. It should be borne in mind, with regard to the observations on waveguide uniformity made in chapter seven, that the waveguide on this lens was formed using an evaporation source situated directly above the lens centre, since only one lens was present on the substrate.

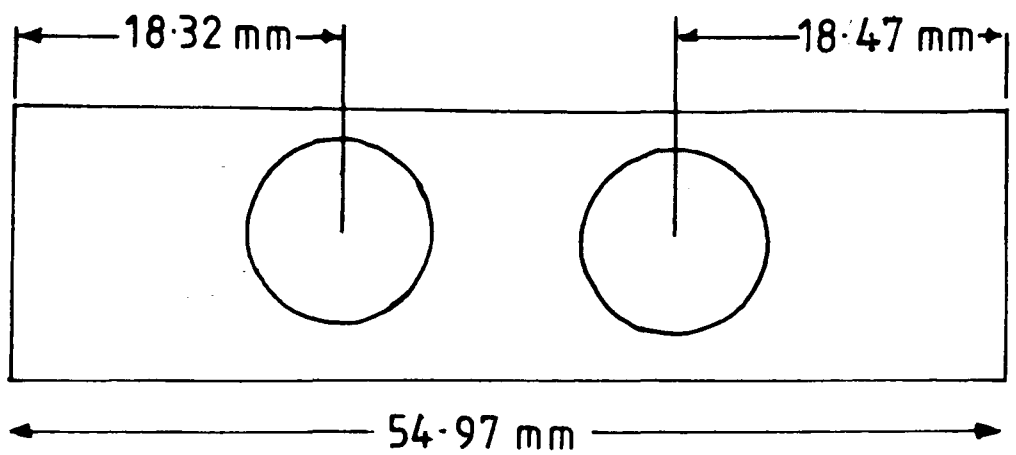
(iv) LENSES GL2/a AND GL2/b

These lenses formed a matching pair, designed for incorporation into an integrated optical spectrum analyzer (IOSA). One lens was intended to collimate the output from a semiconductor laser diode which was to be butt-coupled to the front end of the IOSA substrate. The other lens was intended to act as the Fourier-transforming element in the IOSA. The lens specifications were as for GL1. Here it should be borne in mind that the evaporation source for fabricating the waveguide was NOT held directly above either lens and was instead situated at a point midway between the two lens centres. The dimensions of the lens and substrate are shown in Figure [8.2] . The edges were polished to allow edge-coupling of light into the waveguide. The polishing process reduced the distances from the centre of each lens to slightly less than the design focal length.

8.3 Special tests and techniques

Many techniques are available in order to make objective assessments of the quality of fabricated lenses. These help to determine whether a fabricated lens meets design specifications or to establish whether unforeseen limitations exist which restrict performance.

Four categories of test have been implemented in the course of the present study. These range from purely qualitative assessments of images produced by the lenses, through quasi-ray methods designed to measure aberrations, to direct quantitative measurements taken from images recorded using photo-metric



Lens : 'SOTTINI' design

$f_1 = \infty$

f_2 = 18·5 mm

outer diameter = 10·0 mm

inner region diameter = 7·4 mm

FIGURE [8·2] DIMENSIONS OF GL 2 SUBSTRATE AFTER POLISHING OF THE ENDS

techniques. One technique that was not used was that of interferometry, although it may be noted in passing that this is the only means of entirely separating the geometrical and interference/diffraction aspects of lens imagery [1]. Interferometry, however, is a difficult technique to apply to integrated optical lenses.

A feature common to all the tests reported herein is that a point-source situated at infinity and lying on the optical axis is used throughout. The incident phase-front impinging upon the lens is, therefore, essentially plane (though its amplitude distribution, in keeping with the usual situation for laser-beam illumination, is more nearly Gaussian).

8.3.1 Foucault-, or 'knife-edge' testing

In this test, a sharp 'knife-edge' (eg. a razor-blade) is placed in a given image plane close to the expected focus. The edge is then drawn across the image in a direction perpendicular to the optical axis. Figure [8.3] is a schematic showing the method of implementation of the test. The diagrams on the left-hand side of the figure show knife-edges introduced at selected points in the image space. These serve to block off some of the incident rays whilst allowing other rays to pass. The scene is viewed by an observer situated to the right of the knife-edge.

The effect of the knife-edge on the observed images is seen in the diagrams on the right of Figure [8.3]. These diagrams represent images seen with conventional bulk lenses having circular symmetry, for purposes of clarity. Integrated optical lenses produce only the image observed through the central vertical azimuth.

Figure [8.3] (b) is the most revealing and informative section. In this area of the image the rays emanating from the lens have begun to intersect each other. They intersect at various discrete points for a lens producing aberrations, whereas they intersect at only one point for an unaberrated image, within

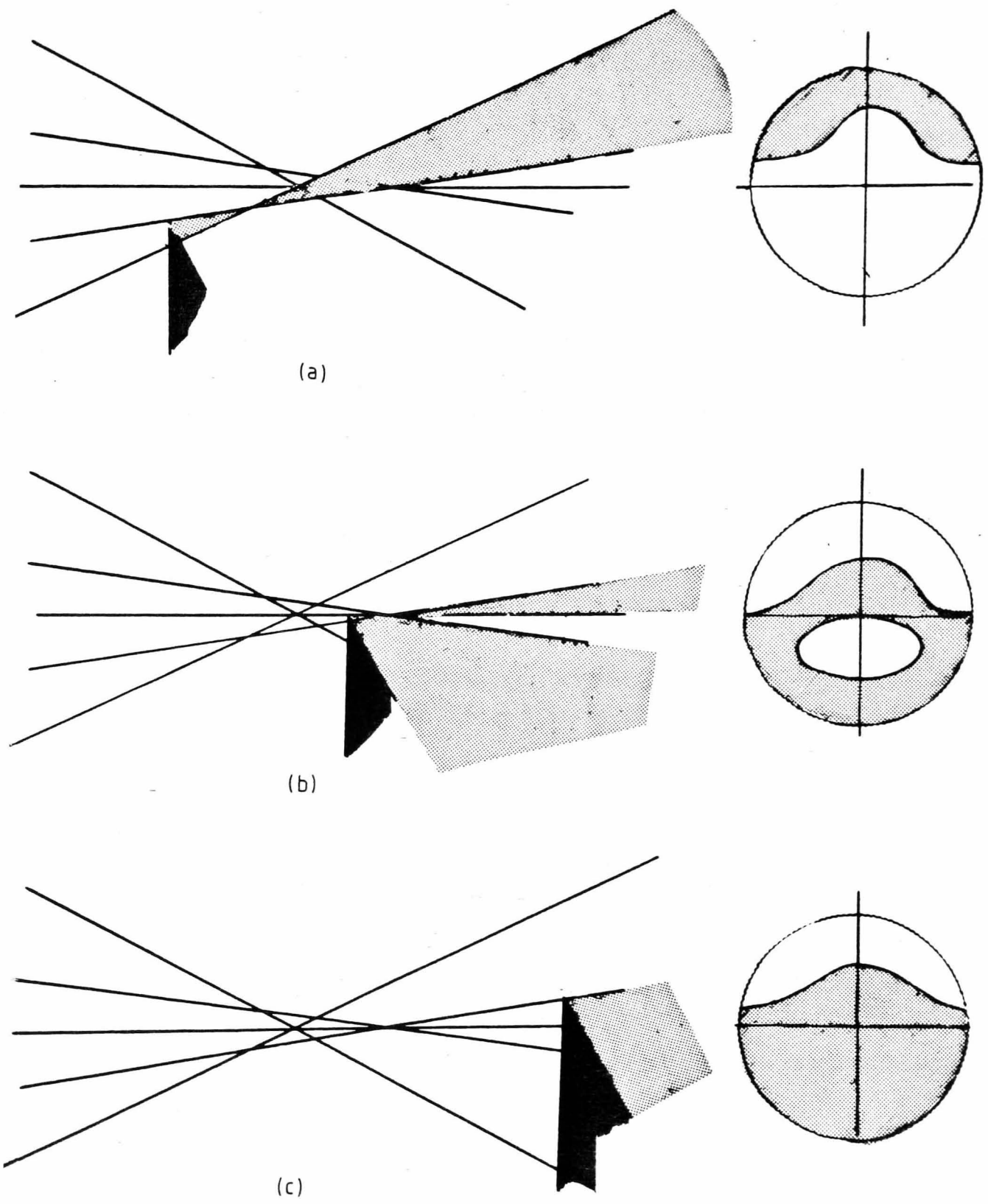


FIGURE [8,3] THE FOUCAULT TEST

the geometrical optics approximation. An aberration-producing lens will form two distinct bright areas separated by a shadowed region. A perfect lens however will cause a smooth continuous transition between shadowed and unshadowed regions to occur as the knife-edge is drawn across the image.

The technique just described is ideal for locating the best point of focus, if the knife-edge is mounted on accurate horizontal and vertical translation stages. The more compact the light distribution is in the image-plane, the more rapidly will the shadow move across the pupil. Under-corrected and over-corrected lenses exhibit characteristic patterns under the conditions of knife-edge testing. Furthermore the technique is highly sensitive to the degree of aberration produced by different zones across a lens aperture. The method represents a simple way of relating the spread of the light in the image plane to the zone of the lens from which the aberrations, if any, arise.

In many cases of interest in integrated optics a re-imaging lens has to be used to gain access to the image produced by the test lens. In such cases, the light has to depart from the surface waveguide. The small transverse dimension of the waveguide then causes considerable beam spreading due to diffraction. Difficulties may consequently arise in correctly interpreting the patterns produced by the moving knife-edge.

In the experiments described here knife-edge testing was largely used to provide a rapid, approximate evaluation of lens quality. If necessary, however, the method can be used to yield quantitative information so precisely that image errors as small as one-tenth of a wavelength can be detected [1].

8.3.2 The Hartmann Test

The geometric aberrations of a ray can be measured directly by carrying out an experimental ray-trace. Tests which involve measurements on experimentally produced 'rays' are known as

Hartmann tests [1]. A schematic of the method is shown in Figure [8.4]. A broad diameter optical beam impinges on a diaphragm made of glass or some other material, which is opaque except for several circular apertures of small diameter. These apertures serve to convert the beam into a series of thin pencils which simulate light rays (light rays are of course, conceptually, infinitely narrow). The diaphragm is placed as close to the aperture of the test lens as possible to minimise unwanted diffraction effects. The rays travel through the test lens and intercept each other in the image space at points determined by the aberrations produced by the zone of the lens aperture through which the rays pass.

Photographs taken at two known image planes can yield the ray-intercept coordinates at these planes. The ray positions in the whole of the image space are then completely determined, if the space is homogeneous so that the rays follow straight lines. Photo-detector arrays [2] or calibrated microscopes [3] may be used to measure the aberrations instead of photographic plates. Single rays are not used due to the difficulty in obtaining a reliable reference position. Furthermore a single ray will in general give rise to an Airy diffraction disk in the given measurement plane which forms a diffuse patch of light a few times bigger than the aberration component itself. Complimentary pairs of rays are used instead, which enable the interception points of the two rays to be measured with a great deal of precision due to the formation of an interference pattern with a bright central spot. With photographic plates, the data resulting from the test may be examined at leisure, and thus multiple ray diaphragms may be utilised. When detector arrays or calibrated microscopes are used, however, the data points are recorded as the test is carried out and single pairs of rays should be used to avoid confusion.

The schematic of Figure [8.4] shows a set-up for detecting directly or indirectly both components of spherical aberration. Where the focal region is accessible with precision, the image sensor may be used to measure TA, the lateral component and LA,

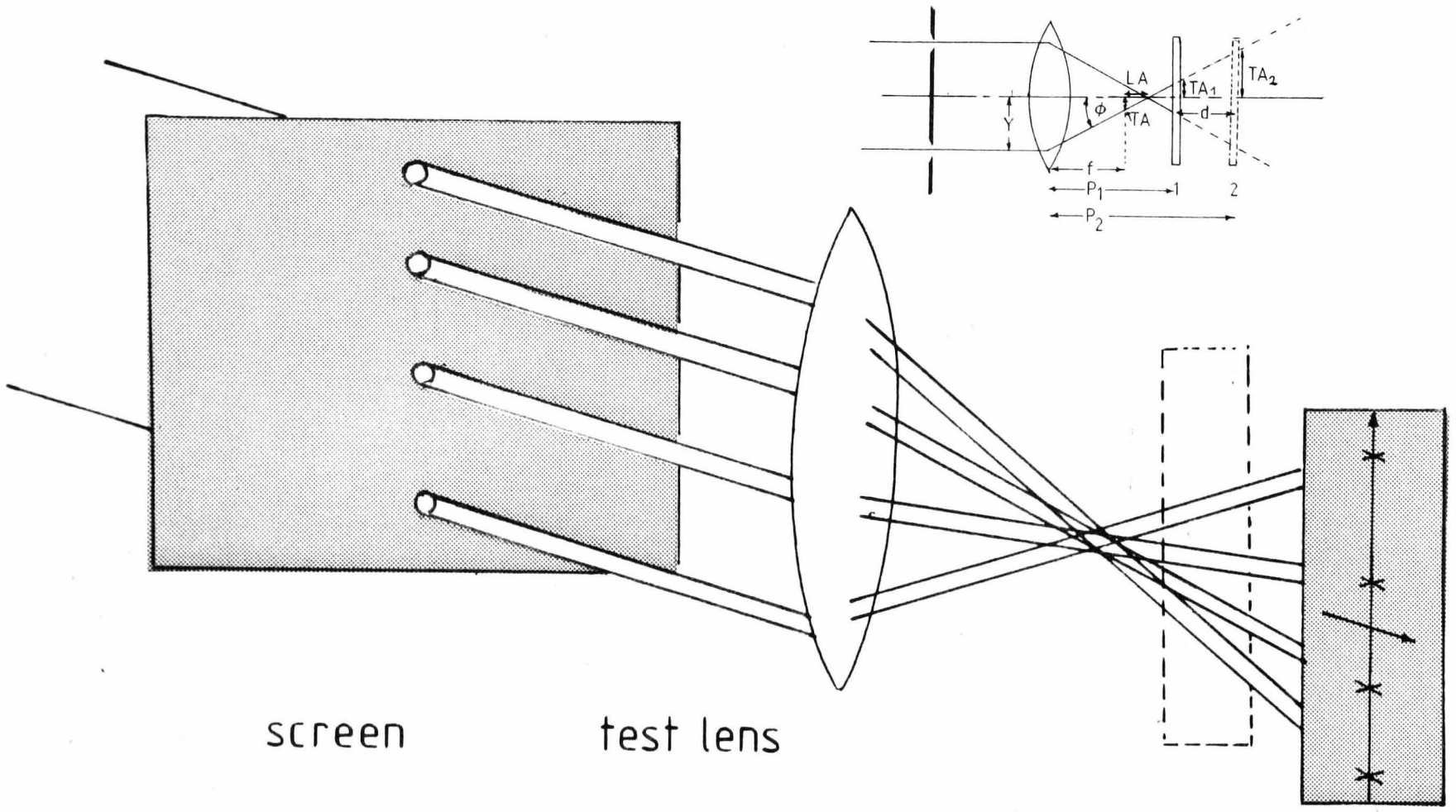


FIGURE [8.4]

HARTMANN TEST

the longitudinal component directly. More commonly in integrated optics, the focal region is not accessible and measurements must be carried out in two planes distant from the paraxial image plane. A high-quality microscope objective may be used to magnify the distances involved. It is only slightly more complicated to devise a set-up for measuring off-axial aberrations such as coma.

Planes 1 and 2 shown in the Figure are the planes at which measurements are taken. The distances P_1 and P_2 of the planes from the centre of the lens are presumed to be known. The spacing between the two planes is d and the measured ray intercepts at the two planes are given by TA_1 and TA_2. The paraxial focal length of the lens is assumed from design calculations to be f. (If the paraxial focal length turns out not to be f, this will be detected by the aberration curves). The required components of spherical aberration are TA and LA. ϕ is the angle which the given ray makes with the optic axis in the image space, and Y is the ray entrance height. The signs of the quantities as shown in the diagram are: LA, TA_1, TA_2, f, P_1, P_2, d all positive; Y, TA, ϕ all negative. From simple geometry:

$$\frac{(TA_2-TA_1)}{d} = \frac{(TA_1-TA)}{(P_1-f)} = \frac{TA}{LA} = \tan\phi \qquad (8.2)$$

Hence:

$$TA = TA_1 - \frac{(P_1-f)(TA_2-TA_1)}{d} \qquad (8.3)$$

and:

$$LA = \frac{d\ TA}{(TA_2-TA_1)} \qquad (8.4)$$

The system, if used with care, can detect aberrations of tens of microns and can locate paraxial foci with an accuracy of 5 to 10 microns. Aberrations an order of magnitude above this are easily detected.

Figure [8.5] shows the type of pinhole mask used to create ray pencils in the experiments reported here. Ten different hole pairs with spacings between 0.25mm and 9.0mm were produced on aluminium coated glass slides using photolithographic techniques to ensure a high degree of accuracy. The hole diameters used were 0.125mm and 0.25mm and this was found to give sufficiently high ray brightness and entrance pupil placement accuracy, together with low far-field diffraction, all factors which need to be balanced in order to optimise test accuracy. The wider diameter holes were used to increase the light throughput in overlay Luneburg lenses where the lens material, as reported in chapter six, is quite lossy.

8.3.3 Resolution test using the interaction between light and sound

It is well-known that an acoustic wave can interact with an optical wave so that optical radiation may be diffracted into one or more subsidiary orders which are frequency-, irradiance-, and direction-dependent on quantities associated with the acoustic wave [4,5]. The acoustic wave causes a perturbation in the density of the material medium supporting it which, in turn, perturbs the refractive index of the material. Consequently a grating is set up which deflects part of the incident optical radiation depending upon, among other things, the angle of incidence of the optical beam. If the angle between the optical and acoustic beams is Θ, then constructive interference occurs for angles satisfying the condition:

$$\sin \Theta_B = \frac{\lambda}{2n\Lambda} \qquad (8.5)$$

where λ is the wavelength of the optical beam in vacuum, Λ is the wavelength of the acoustic wave in the medium and n is the unperturbed value of the refractive index. Θ_B is known as the Bragg angle and equation (8.5) is known as the Bragg condition.

It transpires that surface acoustic waves (SAW), ie acoustic

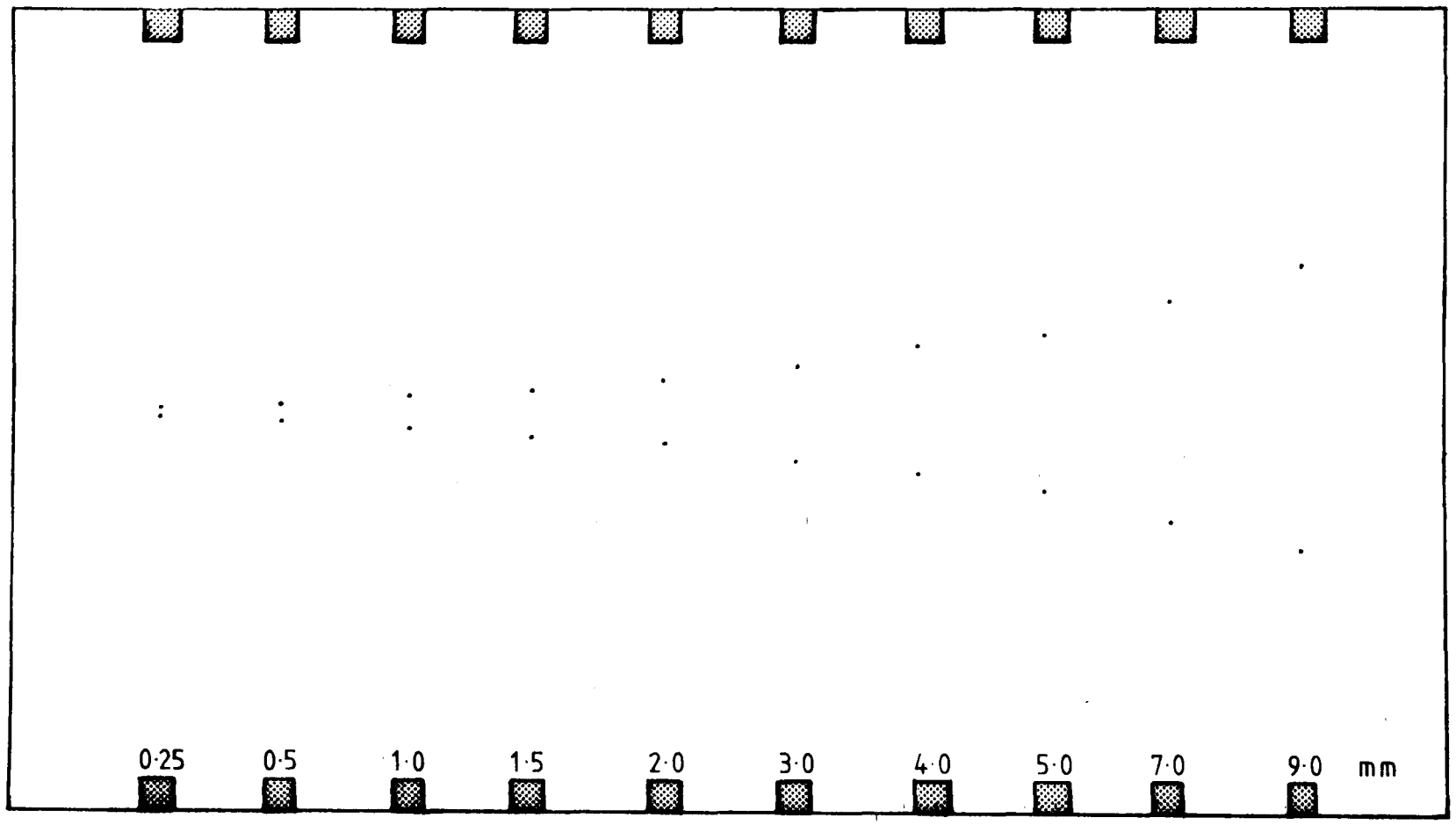


FIGURE [8,5] PINHOLE MASK FOR HARTMANN TEST
pinhole diameters: 0·125 / 0·25 mm

waves which are constrained to travel at the surface of a medium, can deflect surface-guided light very efficiently, due to long interaction lengths between the propagating optical and acoustic wave-fields, and the acoustic power density implied by a surface-confined wave. This fact is exploited in the integrated optical spectrum analyzer. The interaction between the two types of surface wave can be used to test the resolving power of a lens.

Surface acoustic waves are closely approximated for the purposes of the spectrum analyzer by uniformly weighted plane waves within a diffraction angle $\delta\phi$ [6]. Two acoustic waves differing very slightly in frequency then approximate to two closely spaced point sources giving rise to optical waves impinging upon the lens. The deviation in angle between the two waves, each satisfying the Bragg condition, is $\delta\Theta$, where:

$$\delta\Theta = \frac{\lambda \delta f_s}{nv} \qquad (8.6)$$

and where δf_s is the difference in frequency between the two sound waves and v is the sound velocity, which is assumed constant for both waves. The diffracted optical waves may be considered to have constant amplitude over an aperture width D. In order to obtain separation between the two waves through the mechanism of Fraunhofer diffraction over a reasonable substrate distance, a lens (an element which causes contraction of the distances over which Fraunhofer diffraction takes place [7]) is used. The displacement between the beam-spots in the focal plane is given by:

$$\delta y = \frac{f \lambda \delta f_s}{nv} \qquad (8.7)$$

where f is the focal length of the lens. The minimum frequency interval that can be resolved may be estimated using the Rayleigh criterion [8]. It turns out to be of the order of:

$$f_{smin} = v/D = 1/T \qquad (8.8)$$

T as defined by equation (8.8) is the 'access-time' of the device, which may be interpreted as the time which the device takes to acquire a new frequency slot across its entire aperture, ie, with maximum resolution. A real lens will be able to resolve a frequency interval δf_{sreal} where $\delta f_{sreal} \geqslant \delta f_{smin}$. An experimental measure of the smallest resolvable frequency interval therefore serves to indicate how closely the lens comes to attaining optimal performance.

8.3.4 Direct observation of the image-space irradiance profiles

The measurement of light is the domain of a branch of optics called photometry. Detectors of light do not respond directly to the electric- or magnetic-field amplitude of an incident optical wave-field. Instead, they are sensitive to quantities such as the radiant energy density or the time rate of flow of radiant energy. The time rate of flow of radiant energy can be interpreted as the power associated with a wave-field, measured in watts. The rate per unit area at which radiant energy arrives at a surface for a wave-field impinging at normal incidence is the power density, a quantity called irradiance in optics [7]. The average rate per unit area at which energy is carried in the direction of propagation is, however, commonly called intensity [8,9]. In recent times, a minor dispute has arisen in the optical literature as to which term is more appropriate when describing measured quantities, but irradiance will be the term used in the present work.

It is the variation of this quantity over a region in space, ie the irradiance distribution, that is measured when photo-electric detectors are used. The irradiance is given by the squared modulus of the complex amplitude of the wave-field and thus irradiance distribution measurements provide detailed information on the nature of the wave-field. In measurements of wave-fields in the image-space of a lens, the height and width of the central blur spot and any associated sidelobes are readily

ascertainable, as is the presence and quantity of background scattering. Photo-electric methods can be used to provide electrical output linearly related to irradiance distributions over a range of greater than 10000 to 1, or 40dB.

The irradiance distribution can be used to evaluate the energy encompassed over increasing distances from the centre of an image. Encompassed energy curves were introduced in chapter five, and are more commonly known in the bulk optics of rotationally symmetric media as encircled energy curves. Encompassed energy curves are often the most useful and easily interpreted quality criteria of a fabricated lens.

Irradiance distribution measurements were the tests most used to evaluate lens performance in the study reported herein.

8.4 Comparisons between experiment and theory

It is well-known that the irradiance pattern at the focal plane of an ideal lens can be regarded as a scaled and shifted measure of the angular-spectral distribution of the transmittance function at the entrance-pupil of the lens [7] (strictly speaking the relevant complex amplitude quantities should be considered; however photo-detectors are insensitive to phase and the pattern can only be described in terms of real, measurable quantities such as irradiance). To a first approximation, the observed patterns reported here depend only on the transmittance function of the limiting stop in the system, which may be the aperture of the lens itself and which is a purely real rectangular function, and the amplitude variation of the incident beam, which is usually considered Gaussian. In reality, the measured irradiance patterns are dependent on many factors including the aberrations produced by the test lens, the variable transmittance function of the input prism and the test lens itself, and the properties of the re-imaging instruments. The net result is a complicated multiple-convolution integral. The truncated Gaussian incident-field approximation may, however, be used to calculate the diffraction pattern in the ideal case, and the closeness of

approach that the real lens makes to this ideal may be taken as a measure of its quality.

The Gaussian variation in the amplitude of the input optical beam may be described by: [10]

$$A(x) = \exp(-4x^2/a^2) \qquad (8.9)$$

where the peak amplitude is here normalized to 1, and x is the transverse coordinate at the input plane. The phase-front is considered to be plane at the entrance-pupil of the lens, ie the incident beam waist occurs at the lens entrance. The parameter a is the 1/e amplitude beam-diameter. Since the irradiance is given by:

$$I(x) = |A|^2 \qquad (8.10)$$

the parameter "a" represents the $1/e^2$ diameter of the irradiance profile. By taking the one-dimensional Fourier Transform of the input amplitude function, the optical amplitude distribution in the back-focal-plane of an ideal lens is obtained. In consequence:

$$I(x_1) = \left| \int_{-\infty}^{\infty} \exp(-4x^2/a^2)\exp\{-j[(kx_1)/f]x\}dx \right|^2$$
$$= C \exp\{-2(ak/4f)^2 x_1^2\} \qquad (8.11)$$

where f is the focal length of the lens, k is the free-space optical propagation constant and x_1 is the transverse coordinate in the focal plane. C is a constant related to the total input power.

Equation (8.11) leads to a simple relationship between the beam-diameters of the Gaussian irradiance distributions at the input plane and the focal plane, a and a' respectively:

$$a' = \frac{4f\lambda_o}{\pi na} \qquad (8.12)$$

where λ_o is the wavelength of light in free space, and n is the effective refractive index which the light encounters in the waveguide.

Marom, Chen and Ramer [11] have calculated the effects of beam truncation by an aperture with a rectangular transmission function. For aperture diameters smaller than the untruncated input beam diameter, the values given by expression (8.12) can underestimate the diffraction-limited beam diameter at the focus. It is a simple matter to estimate the diffraction-limit in the case of uniform illumination of a rectangular aperture of width b. A sinc^2 irradiance pattern is observed at the focal plane in such a case:

$$I(x_1) = \left[\frac{\sin\left(\frac{\pi n b x_1}{f\lambda_o}\right)}{\frac{\pi n b x_1}{f\lambda_o}} \right]^2 \qquad (8.13)$$

The function given by equation (8.13) is a one-dimensional version of the Airy disc function. The $1/e^2$-diameter, b', of the central lobe of this function is related to the aperture diameter b by:

$$b' = \frac{4.4 f \lambda_o}{\pi n b} \qquad (8.14)$$

an expression very similar to the one obtained for untruncated Gaussian beams, equation (8.12). The spacing d' between the nulls of the function is given by:

$$d' = \frac{2f\lambda_o}{nb} \qquad (8.15)$$

Equation (8.11) should describe diffraction-limited performance when truncation occurs at diameters greater than the $1/e^2$ irradiance diameter, whereas, for truncation at diameters

less than the $1/e^2$ diameter, equation (8.13) should predict the diffraction-limited performance.

8.5 Apparatus and experimental procedures

The experimental set-up used to investigate the distribution of light in the image-space is shown in Figure [8.6]. The set-up differed only slightly for the various other tests.

A He-Ne gas laser was used as the light-source in all the experiments. The laser operated at a wavelength of 0.6328um and had a power output, typically, of 5mW. In order to reduce the chances of optical damage in the Ti-$LiNbO_3$ waveguide/substrate system, a variable optical attenuator was used to reduce the output power. The light was chopped mechanically at a frequency of 1kHz before being expanded using an f/4 spatial filter and a collimator usually set to give a magnification of 12. Figure [8.7] shows the expanded beam profile. The $1/e^2$ irradiance points were separated by 7.2 mm. The lenses were investigated over a wide range of input beam-widths, the variation in width being accomplished by a variable rectangular stop. It was not feasible to vary the degree of expansion and collimation continuously such that approximately Gaussian profiles could be obtained throughout. The stop was placed as close as possible to the input coupling-prism. The size of the aperture created by the stop, and therefore the effective diameter of the beam entering the test lens, was measured, using a measuring microscope, to an accuracy of 10.0 μm. In Hartmann tests, the mask used to generate narrow pencil-beams approximating to rays was interposed instead of the stop.

The coupling prism was made of rutile, a uniaxial crystal of high refractive index. The c-axis of the prism was parallel with the c-axis of the $LiNbO_3$ substrate (which is also a uniaxial crystal) and both were perpendicular to the general direction of optical propagation. The result was that the extraordinary component of refractive index was used in all relevant

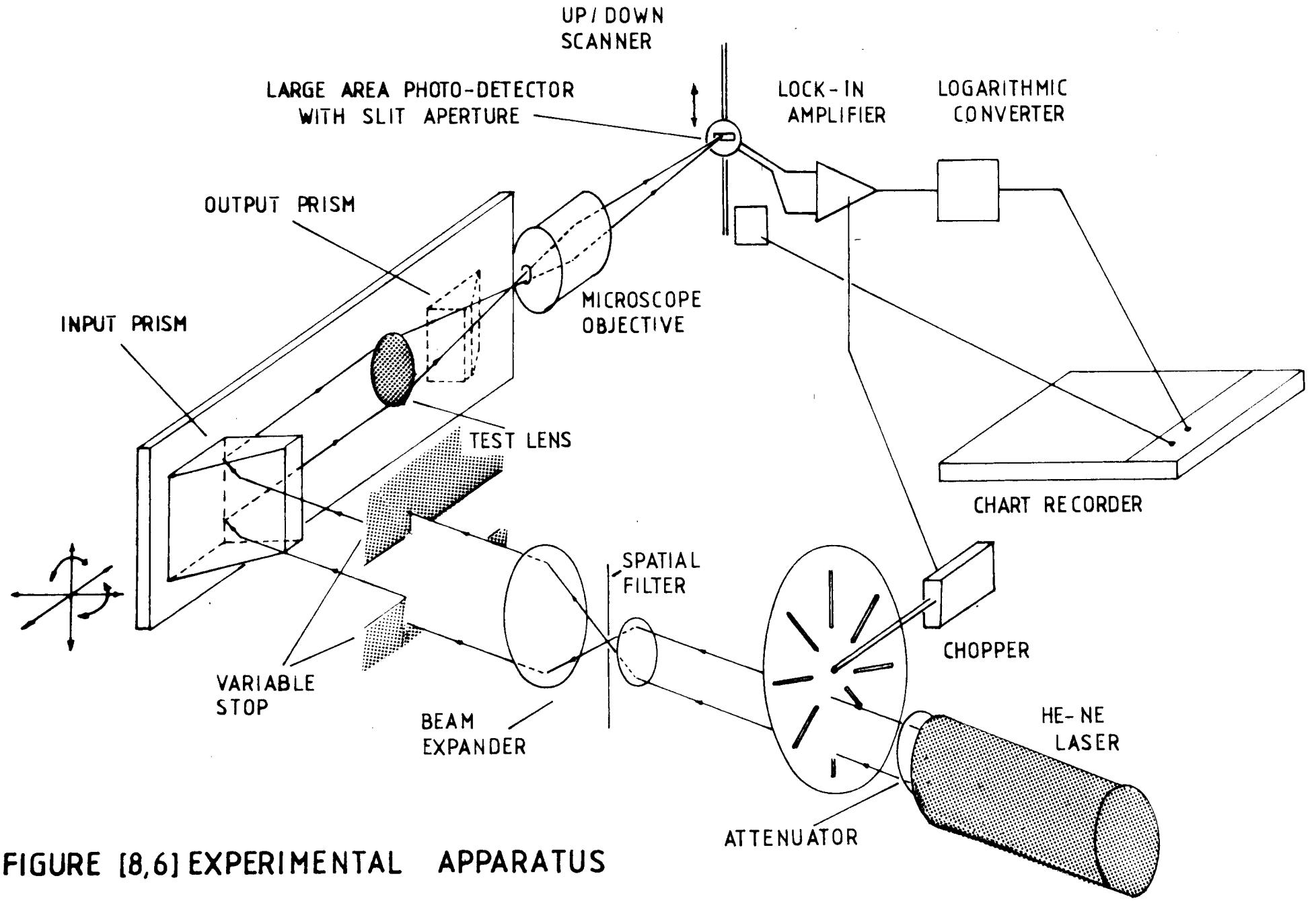


FIGURE [8,6] EXPERIMENTAL APPARATUS

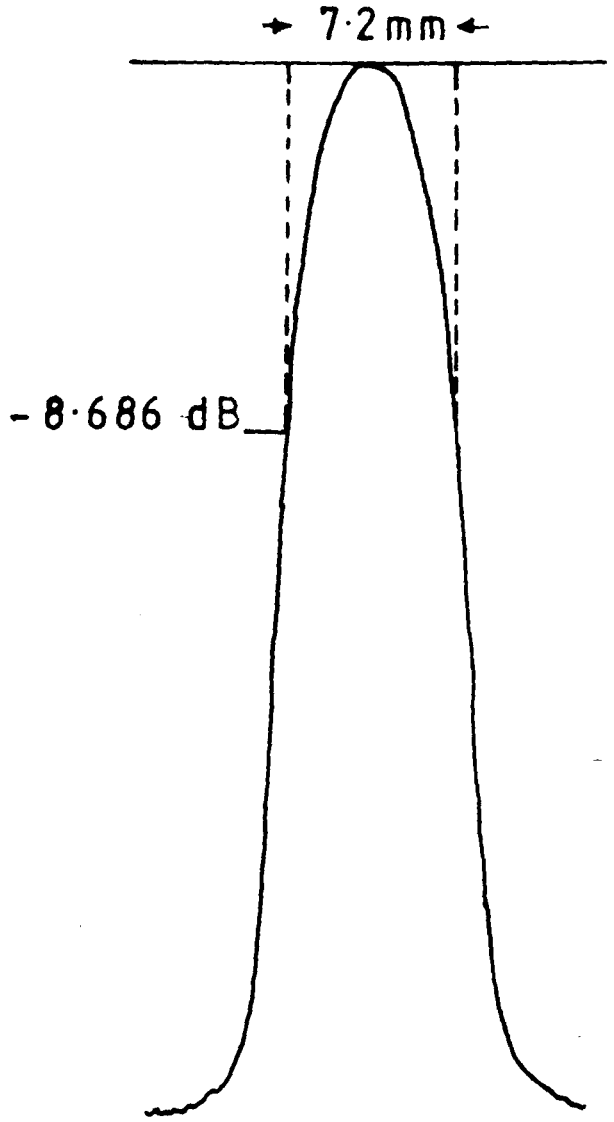


FIGURE [8,7] EXPANDED LASER-BEAM PROFILE

calculations, TE modes only being excited.

For a properly collimated and aligned beam the input prism introduced no aberrations, ie phase-distortions. The prism did, however, affect the transmitted amplitude of the beam. The coupling efficiency of the prism was not constant across the beam cross-section because of the variable air-gap between prism and waveguide. The variable air-gap arises as a result of the screw end-shape used to ensure firm contact between the prism and the wave-guide. The beam-profile was, in consequence, removed from being a purely truncated-Gaussian shape.

The substrate was placed on a test-jig that allowed continuous, accurate linear translation along three axes and rotational movement around two axes. Such a degree of flexibilty was required to ensure efficient coupling of the appropriate guided-wave modes. A 'Micro-controle' digital read-out was used to measure coupling angles to a precision of one-hundredth of a degree.

After the beam had travelled through the test-lens and had begun to converge, a choice of methods was available with which to couple the light back out of the waveguide. In cases where the focus lay close to, or beyond the edge of, the substrate, the edge was finely-polished and edge-coupling thereby facilitated. High-power microscope objectives did not have sufficient depth of field to 'gain access' to the focal region in cases where the focus lay inside the substrate, far from the edge. The substrates could then have been sawn and re-polished. However such a procedure was considered too risky and inaccurate. Another rutile prism was used instead to couple the light out from the waveguide. The problem with the output prism was that converging rays, in contrast with the parallel rays impinging upon the input prism at normal incidence, were refracted considerably by the several material interfaces involved. Aberrations were therefore introduced which served to distort the beamshape from that actually produced by the lens. These aberrations were not necessarily destructive in terms of the quality of the resultant

image. Over-corrected lenses produce marginal rays that intercepted the optical axis at points beyond the paraxial focus. The aberrations introduced by a prism are of opposite sign to those produced by an over-corrected lens, and thus both sets of aberrations tend to cancel, to a certain extent. For under-corrected lenses, of course, the addition of prism-related aberrations to those already present is destructive. Regardless of these pros and cons, the prism effects were considered undesirable as they reduced the fidelity with which the optical characteristics of the test-lens itself could be measured. In this context, it should be observed that the $LiNbO_3$/air interface at the edge of the substrate also introduces aberrations. These aberrations are more easily allowed for, as there is only one interface to be taken into account.

For Foucault tests a razor-edge was passed through the focal plane, either directly through the focus or through a re-imaged version of it. A Tessar camera lens operating at a relative aperture of f/2.8 with a focal length of 55.0mm was used as the re-imaging lens.

High-power Beck or Olympus microscope objectives were used for observing the light irradiance distributions in the focal plane. Magnification ratios were typically 20x, 40x or 100x. A 12x calibrated eyepiece was used to measure ray aberrations and image widths. More usually, the image was projected at long back-focal-plane distances to yield magnification factors of between 160x and 300x. The magnified images were then scanned using a large-area photo-detector, operating in the photo-conductive (linear) mode, which was masked by a 40μm slit. At 160x magnification a slit of these dimensions is theoretically capable of resolving 2 lines/μm. A motor-driven, variable-speed scanning system built in the departmental workshop was used to scan the beam.

The detected signal was amplified by a narrow-band amplifier locked to the frequency of the chopper. The signal was then converted to give a high-resolution logarithmic output, referred

to an electrical reference signal generated by the amplifier. The linearity of the electrical signal of the detector versus the incident optical signal was checked using Neutral Density filters. A linear range of over 20 dB was available.

A Hamamatsu digital image acquisition system was used instead of the custom-built scanner in some experiments. This system consisted of a C1000 camera head and control unit. The camera head incorporated a high-resolution , high-linearity Newvicon N4076 vidicon. A Nikon Micro-Nikkor f/2.8 camera lens of focal length 55mm was used. The camera control unit incorporated line-scan and frame-scan rates that were compatible with computer equipment. A maximum of 1024x1024 lines was available, with each pixel having 256 possible grey scale levels, corresponding to approximately 22dB of optical irradiance. The equipment was connected to a Hitachi monitor. A useful characteristic of this system was the availability of a single-line intensity display which greatly facilitated the location of the focal waists and the points of maximum irradiance. The data acquired by the Hamamatsu system was down-loaded either directly onto an x-y chart recorder or sent to a DEC PDP-11 mini-computer for more comprehensive analysis.

A graph of measured video signal intensity against incident optical irradiance for the Newvicon vidicon is shown in Figure [8.8]. The relationship is extremely linear with a gamma factor of nearly unity, but a dc offset is present.

Barr and Stroud Neutral Density filters were used throughout these experiments in order to calibrate measurements accurately.

8.6 A note on results and discussion

Contrary to usual practice, discussion of these results will follow each sub-section, since it would be impractical to discuss all the results obtained for all four lens-systems together.

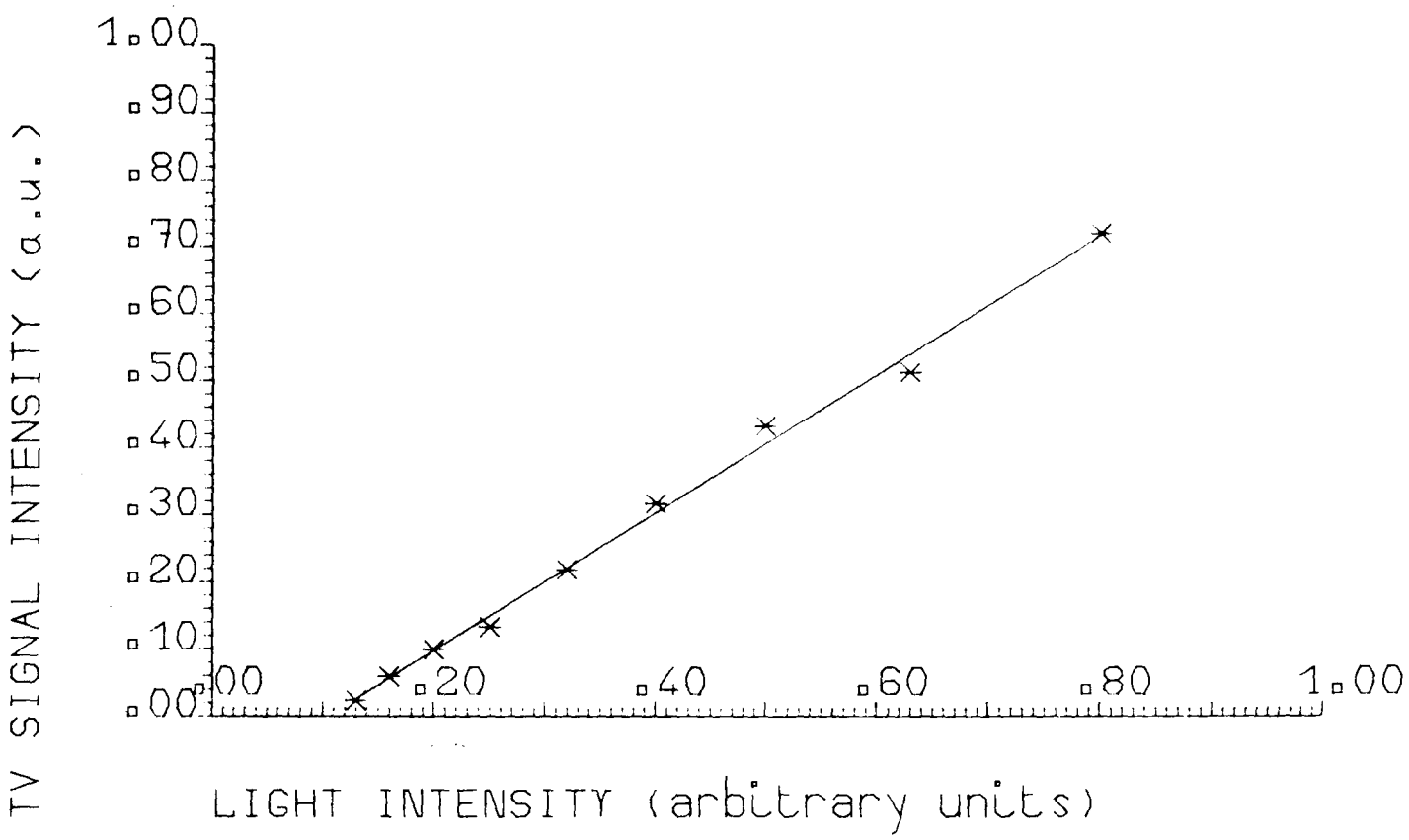


FIGURE [8,8] Relationship between video signal and incident light intensity

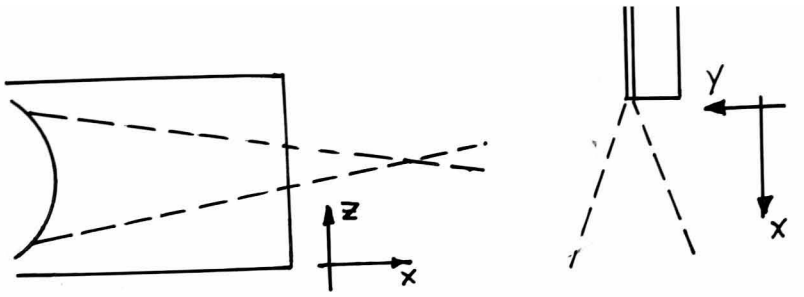
8.7 Results and discussion - lens SL1

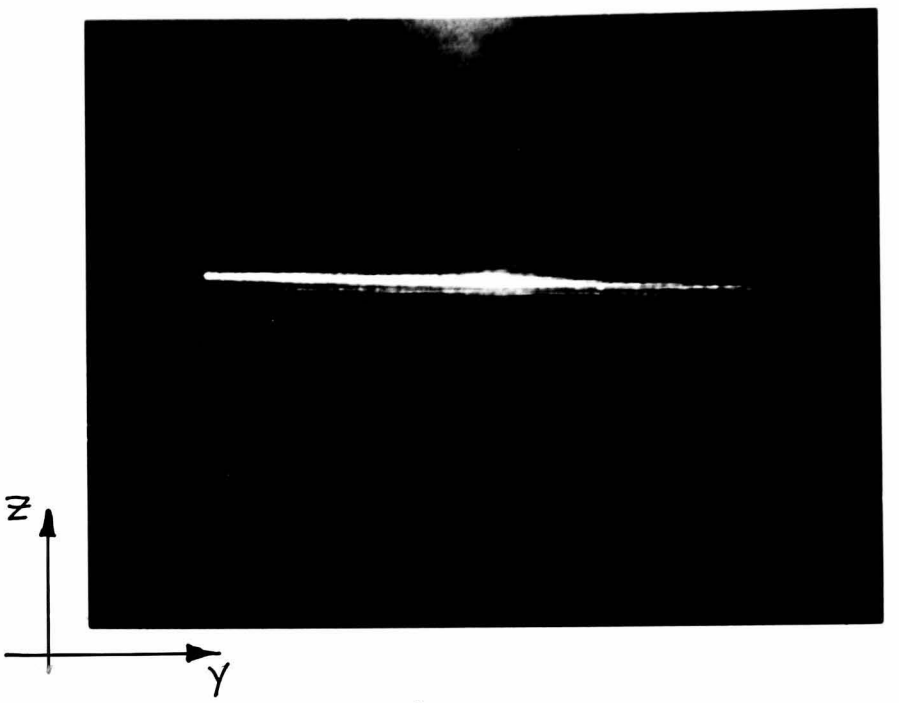
8.7.1 Image irradiance distributions

Focal plane irradiance patterns were recorded for this lens over a wide range of truncated beamwidths from 0.52mm to 7.1mm. These beamwidths correspond to a range of 4.9% to 68% of the maximum available aperture, or to stopped-down f-numbers of between f/100 and f/7.

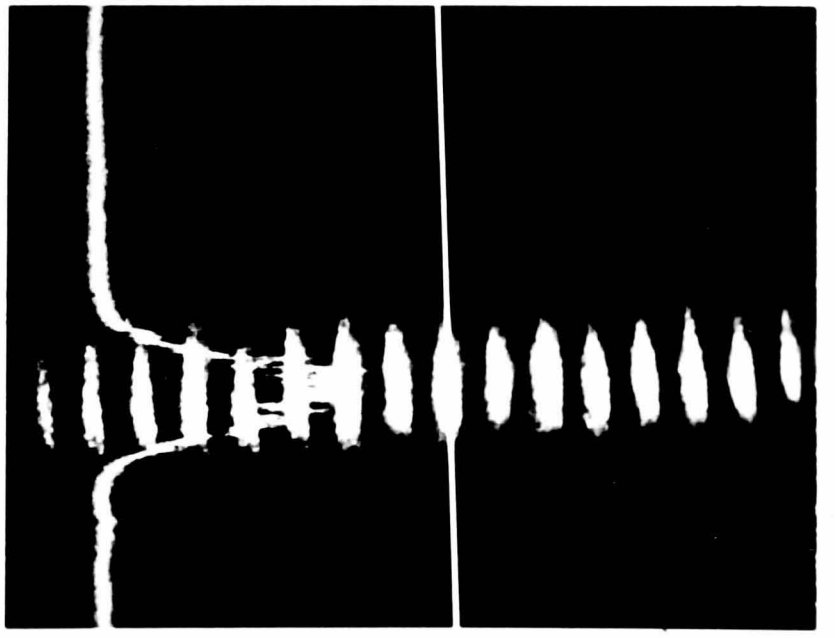
Figure [8.9] (a) shows a typical focal plane image as recorded on a Hitachi TV monitor. The appearance of a broad focal line instead of a small circular blur should not give rise to confusion; it is a consequence of lens SL1 focusing outside the substrate. The small dimension of the waveguide in the x-y plane (perpendicular to the plane of propagation, x-z) caused the light to diverge considerably in this plane. The light levels were still sufficiently high to identify the quantities of real interest, ie the irradiance variation of the light in the z-meridian, corresponding to the vertical direction on the photograph. The situation is analogous to investigating the most pertinent properties of a cylindrical lens in bulk optics. Figure [8.9] (b) shows the light transmitted through a grating of known periodicity, which was used to calibrate the dimensions of the focal image.

Figure [8.10] shows three representative irradiance distributions plotted using a chart-recorder connected to a Hamamatsu image-acquisition system. The truncated beam widths were 0.81mm, 3.65mm and 6.01mm for Figure [8.10] (a), (b) and (c) respectively. The scales in this figure are linear. Theoretical $(\sin x/x)^2$ curves having the same $1/e^2$ spot-diameters as the experimental curves are also shown. The theoretical curves serve to highlight the experimental sidelobe levels, positions and decay rates and degree of sidelobe symmetry. All three experimental curves fit the central lobes of the theoretical curves well. Figure [8.10] (a) shows that the experimental curve





(a)



(b)

FIG [8,9] (a) focal plane of geodesic lens

(b) SAW delay line interposed for calibration

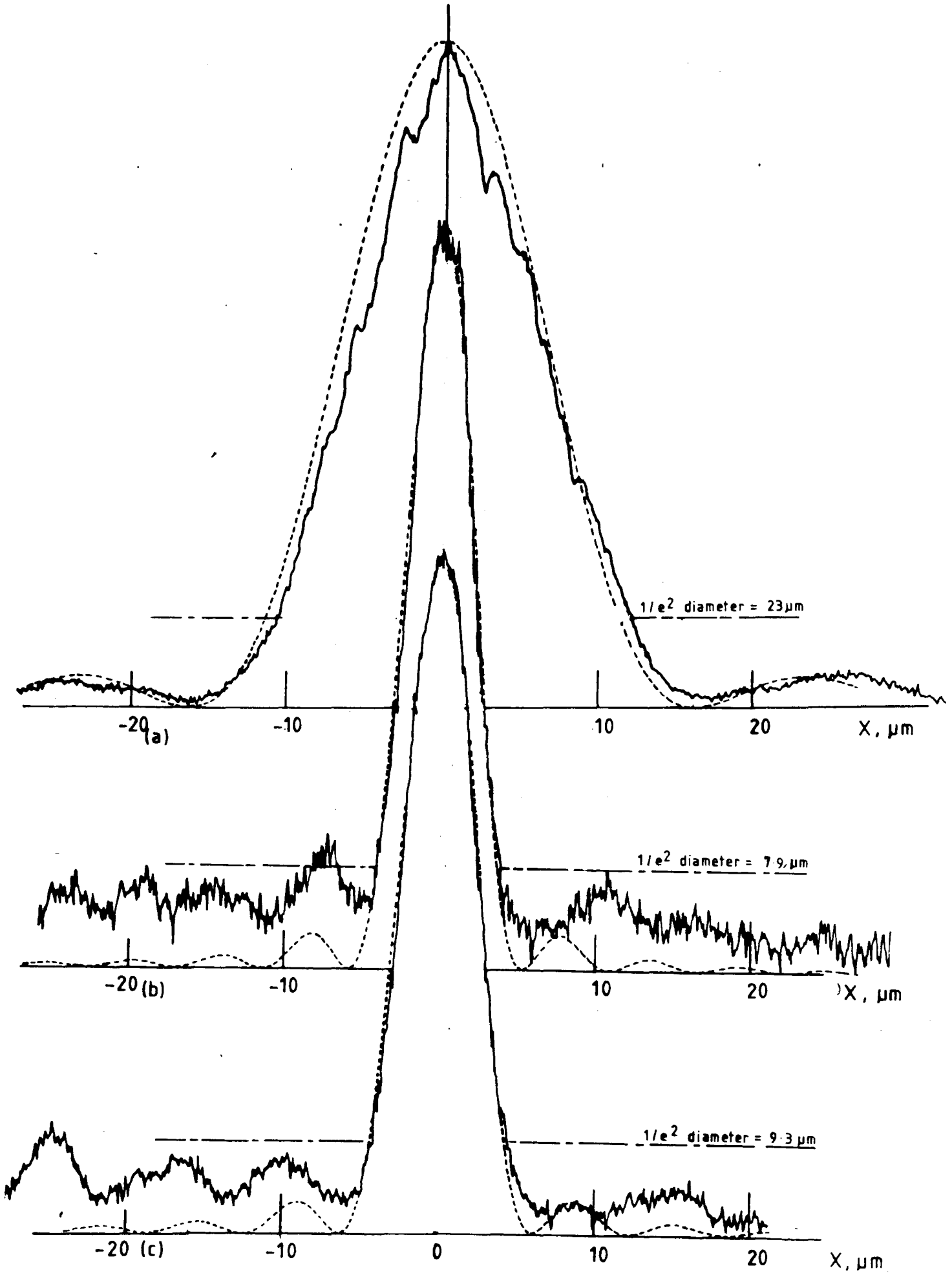


FIGURE [8,10] Irradiance profiles at focal planes of lens SL1 at various values of aperture truncation. $(Sin(x)/x)^2$ curves having same $1/e^2$ widths as experimental profiles are also shown

(a) truncation diameter = 0·81 mm
(b) " " = 3·65 mm
(c) " " = 6·01 mm

fits the theoretical curve well over the entire scanned range, both in terms of sidelobe position and irradiance. The other two experimental curves show sidelobes at a generally higher level than in the theoretical case. There is a hint of a constant dc-type level for which the noise-floor level associated with the dark current of the imaging vidicon may be responsible. These latter curves show increased noise, probably as a result of an increase in scattered light levels associated with larger beam-widths.

8.7.2 Discussion on observed patterns

Some of the sidelobe positions coincide reasonably well with the theoretical positions while others do not. Higher-order experimental sidelobes tend to be further away from the main lobe than expected and a possible reason for this may be the presence of pin-cushion distortion somewhere in the re-imaging process. If the re-imaging system is faithful in reproduction, however, another explanation is available. The presence of irradiance nulls where peaks should occur and vice-versa is a phenomenon known as contrast-reversal. From the point of view of frequency analysis of imaging systems, contrast-reversal is indicative of a negative Optical Transfer Function (OTF) for a lens over a certain band of spatial frequencies [12]. Contrast-reversal is associated purely with the presence of aberrations, notably de-focus and spherical aberrations in the case of symmetrical systems; diffraction-limited systems never display contrast-reversal. Photographs presented in chapter nine of 'Principles of Optics' by Born and Wolf [8] and, strikingly, in the 'Atlas of Optical Phenomena' by Cagnet, Francon and Thrierr [13], show clearly the effect of contrast reversal associated with varying degrees of primary spherical aberration. The higher-order sidelobe irradiance levels can be greater than those of lower-orders if sufficient amounts of aberration are present in the system.

Beyond these qualitative observations on contrast-reversal

as evidence of the presence of spherical wave-aberration, little can be inferred from the irradiance distributions as to the precise shape of the wave-fronts at the exit-pupil of the lens which give rise to the distributions. Partly, this is because of the difficulty (in current practice) of obtaining information on the optical phase. A more fundamental objection exists however. It is certainly true that the wave-front aberration function in the exit-pupil can be used to predict the irradiance distribution at the focal-plane [8]. Such a procedure falls within a domain in physics known, in scattering theory, as the 'direct-source class of problem'. However the retrieval of the wave-front from the irradiance distribution falls within a related though opposite domain, known in scattering theory as the 'inverse-source class of problem'. The very philosophy of this field rests on shaky foundations, as pointed out by Ross et al [14] :

' ... the (inverse-source problem) cannot be solved simply by applying the deductive process in reverse direction: strictly speaking the inverse problem cannot be solved at all.'

Notwithstanding this underlying pessimism, Ross et al go on to expound cheerily on how to '... approximate the truth by stages ...' . Approximating the truth for the type of problem considered here has been carried out by various authors [15,16,17]. These authors developed computational algorithms for retrieving the source wave-front from irradiance distributions in the presence of noise. The algorithms suffer however from problems of numerical instability, lack of convergence and non-uniqueness of solution. No attempt was made to implement them by the present author.

Only bald statements can therefore be made concerning the irradiance distributions shown in Figure [8.10] :

(i) the central lobes look reasonable, in that there is a good correspondence between the experimental and theoretical shapes;

(ii) the side-lobes are high but, given a spherical-geometry lens, perhaps not unreasonably so;

(iii) the evidence of contrast-reversal is indicative of the presence (but not the degree) of spherical aberration and/or defocus.

8.7.3 The variation of focal spot-size with truncation width

The measured 1/e- and $1/e^2$-diameters of the central lobes of the irradiance distributions arising from 21 different truncation widths of the input beam are presented in Table [8.1]. Several of these measurements were made by Mr. A. Hodkin of Marconi Research. An input beam-width of 3.65mm gave rise to the smallest focal spot-sizes, 5.4μm and 7.9μm for the 1/e- and the $1/e^2$-diameters respectively. The latter result is only 1.55 times the diffraction-limit, a reasonably good figure. The relative aperture of the lens at this beam-width is f/13.9.

The results obtained using edge-coupling and prism-coupling are distinguished from each other in the table. The diffraction-limit calculated using equation (8.13) is also shown.

The results are plotted in Figure [8.11] . As the beam-width increases, the measured focal spot-sizes first decrease steadily and then broadly level out. A feature of the graph is that the prism-coupled results do not follow the general trend of the edge-coupled results and give greater estimates for the spot-sizes. The reason is probably that the output prism produces aberrations of its own, as predicted earlier.

The prism-coupled $1/e^2$-diameter results were ignored and a least-squares fit parabola calculated for the remaining points. The parabola, experimental points and the diffraction-limit curve are shown together in Figure [8.12]. The parabola is less than twice the diffraction-limit over almost the entire range. The parabola, diffraction-limit and a gyration-radius curve derived

TABLE [8,1] Experimentally-observed focal spot-sizes as a function of size of (truncated) input beam for lens SL1

E : indicates result obtained using edge-coupling

P : indicates result obtained using prism-coupling

Truncated beam-diameter, mm	Focal spot-diameter, um 1/e	1/e^2		Diffraction-limit, μm (1)
* 0.52	25.4	36.6	E	35.8
0.81	16.3	23.0	E	23.0
* 1.05	16.1	22.0	E	17.7
1.13	15.2	27.6	P	16.5
1.23	14.7	21.0	P	15.1
* 1.27	12.6	19.6	E	14.7
* 1.63	8.3	14.1	E	11.4
1.77	10.2	17.2	P	10.5
1.82	7.8	11.5	P	10.2
* 1.92	8.3	11.4	E	9.7
* 2.19	8.0	11.4	E	8.5
2.49	8.3	13.4	P	7.5
2.6	8.5	17.4	P	7.2
* 2.62	7.7	10.6	E	7.1
* 3.09	6.6	9.3	E	6.0
3.65	5.4	7.9	E	5.1
4.05	5.7	8.1	E	4.6
* 4.45	5.9	9.7	E	4.2
* 5.05	6.3	9.1	E	3.7
6.01	6.5	9.3	E	3.1
* 7.1	6.1	9.1	E	2.6

NOTES: (1) diffraction-limit is calculated for a Gaussian beam

* measurements carried out by Hodkin

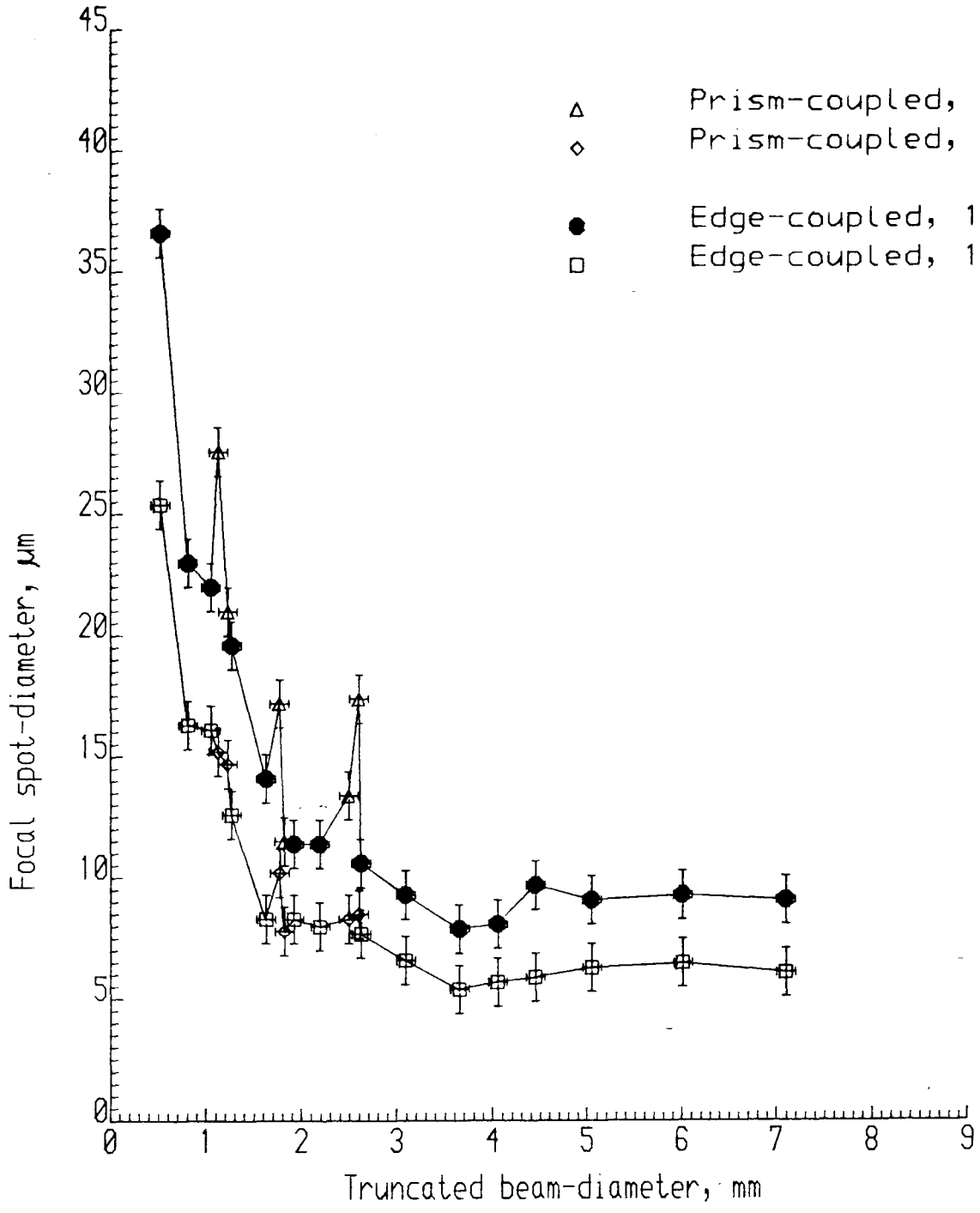


FIGURE **[8,11]** Measured sizes

of image-plane focal spots

of lens SL1

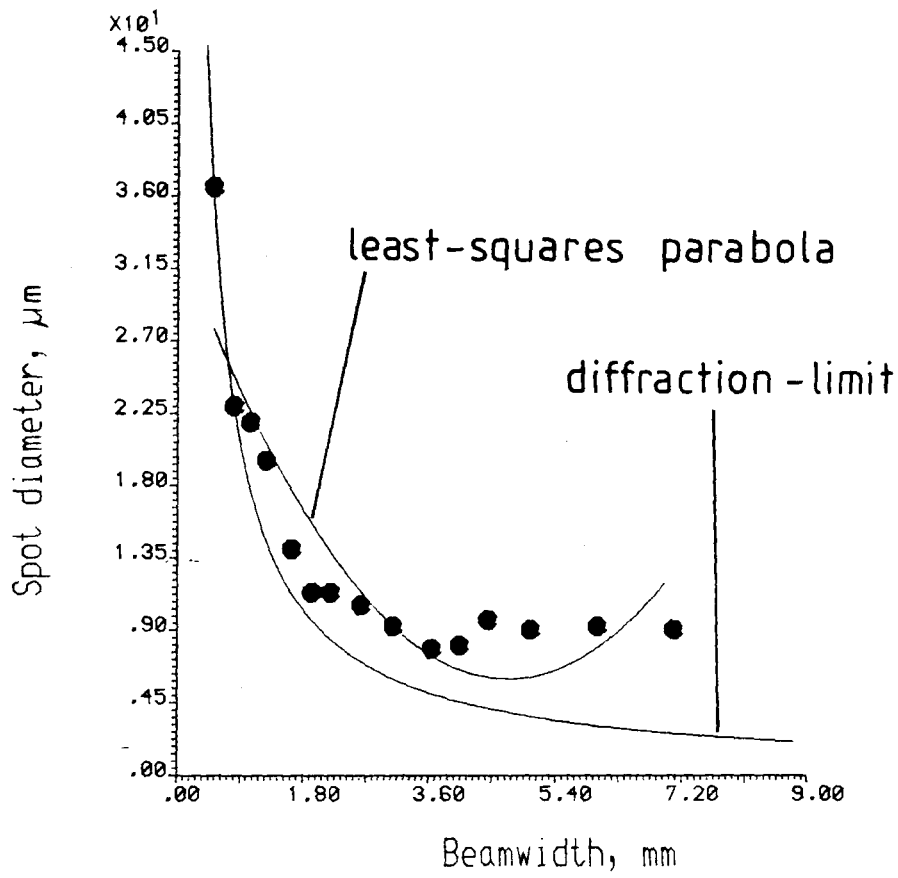


FIGURE [8.12]
SL1 spotsizes and theoretical limit

from raytracing are plotted together in Figure [8.13]. A qualitative reason for choosing a parabolic fit is given below.

8.7.4 Discussion

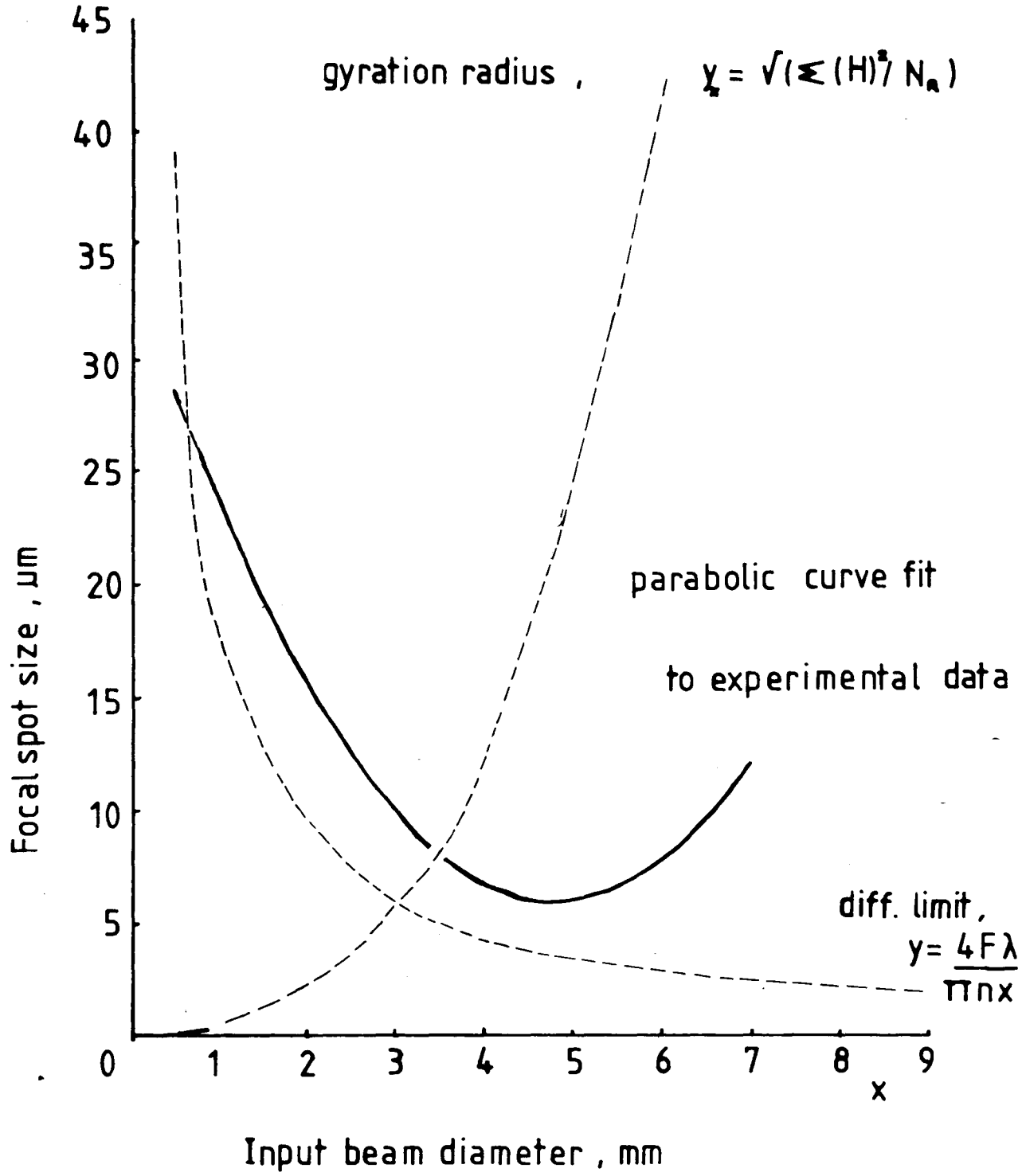
With reference to Figure [8.13], it is apparent from theory that aberration effects are small at small beam-widths and that diffraction effects dominate. Gradually, as the beam-width is increased, aberration effects and the diminishing effects of diffraction tend to balance. At beam-widths approaching the full available aperture, aberration effects dominate. It should be pointed out that no attempt is made in Figure [8.13] to quantify the relative weight of diffraction and aberration effects. For reasons that have long separated physical from geometrical optics, it would be unreasonable to do so. A definite minimum would be predicted, however, as the lens moves from being diffraction-linited to being aberration-limited. Such behaviour was that seen under experimental conditions.

8.7.5 Depth of focus

The evolution of the image-plane irradiance profiles through the focal region is shown in Figure [8.14]. The side-lobe pattern on only one side is presented. The other side of the pattern was not imaged on the screen because of the large magnification used. A truncated input beam-width of 3.65mm was used.

A dip exists in the centre of the irradiance profile recorded at a distance of (f - 235μm) from the lens. The irradiance profile recorded at the other extremity, (f + 275μm) , exhibits a flat topped central lobe. The minima of the profiles are not true nulls. The energy distribution in the side-lobes, as characterised by the relative side-lobe levels, does not fall off in a regular manner with increasing side-lobe order. All these details indicate that the lens at this aperture suffers from spherical aberration.

The information obtained from these profiles is condensed in



FIG[8,13]Theoretical and experimental spot sizes

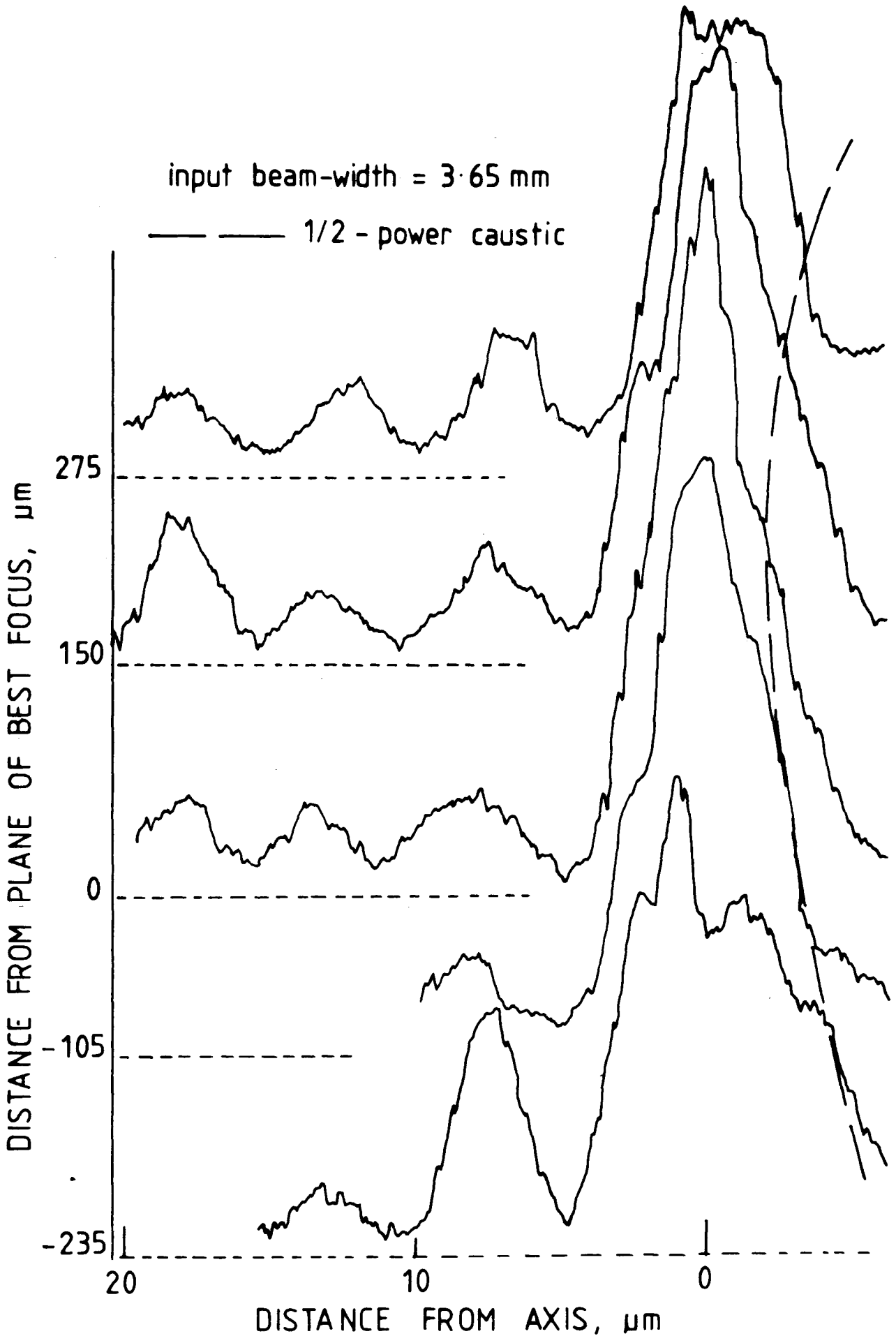


FIGURE [8,14] INTENSITY PROFILES AT SELECTED IMAGE PLANES , LENS SL1

Figure [8.15] . Three curves are shown. The top curve shows the fall in peak irradiance relative to the peak irradiance at the plane of best focus. The middle curve shows the height of the first side-lobe relative to the peak irradiance at each image-plane. The third curve shows the 1/e-diameter at each image-plane, normalized to the 1/e-diameter at the plane of best focus. The asymmetry of the curves is, once again, characteristic of the presence of spherical aberration.

These curves taken together constitute a basis for estimating the depth of focus of the lens at the given aperture, and thus the tolerances which may be established on the setting of the detector plane. The detector plane could be moved 150μm further away from the lens than the best focal plane with only a 1dB drop in peak irradiance, a 1dB increase in nearest side-lobe irradiance and a broadening of the beamwidth by a factor of 1.15.

8.7.6 Direct measurement of spherical aberration

A Hartmann test was conducted on lens SL1 to measure the spherical aberration directly. Both lateral and longitudinal components were measured. The results are shown in Figures [8.16] and [8.17]. The computed curves obtained using Southwell's method (chapter four) are also shown. Theoretical and experimental curves are in excellent agreement.

8.7.7 Acousto-optic resolution tests

The ability of lens SL1 to resolve two parallel beams with slightly different incidence angles was tested using the acousto-optic method. Figure [8.18] shows diode-array images in the focal -plane obtained after acousto-optic diffraction of a 1mm truncated beam, for two values of input rf-signal electrical power. A re-imaging lens was used throughout the acousto-optic tests to enable the diode-array to have easy access to the focal plane. Using ND filters to attenuate the optical beam by fixed amounts, the diffracted light signal falling on a large-area photo-detector was measured and compared with rf-signal

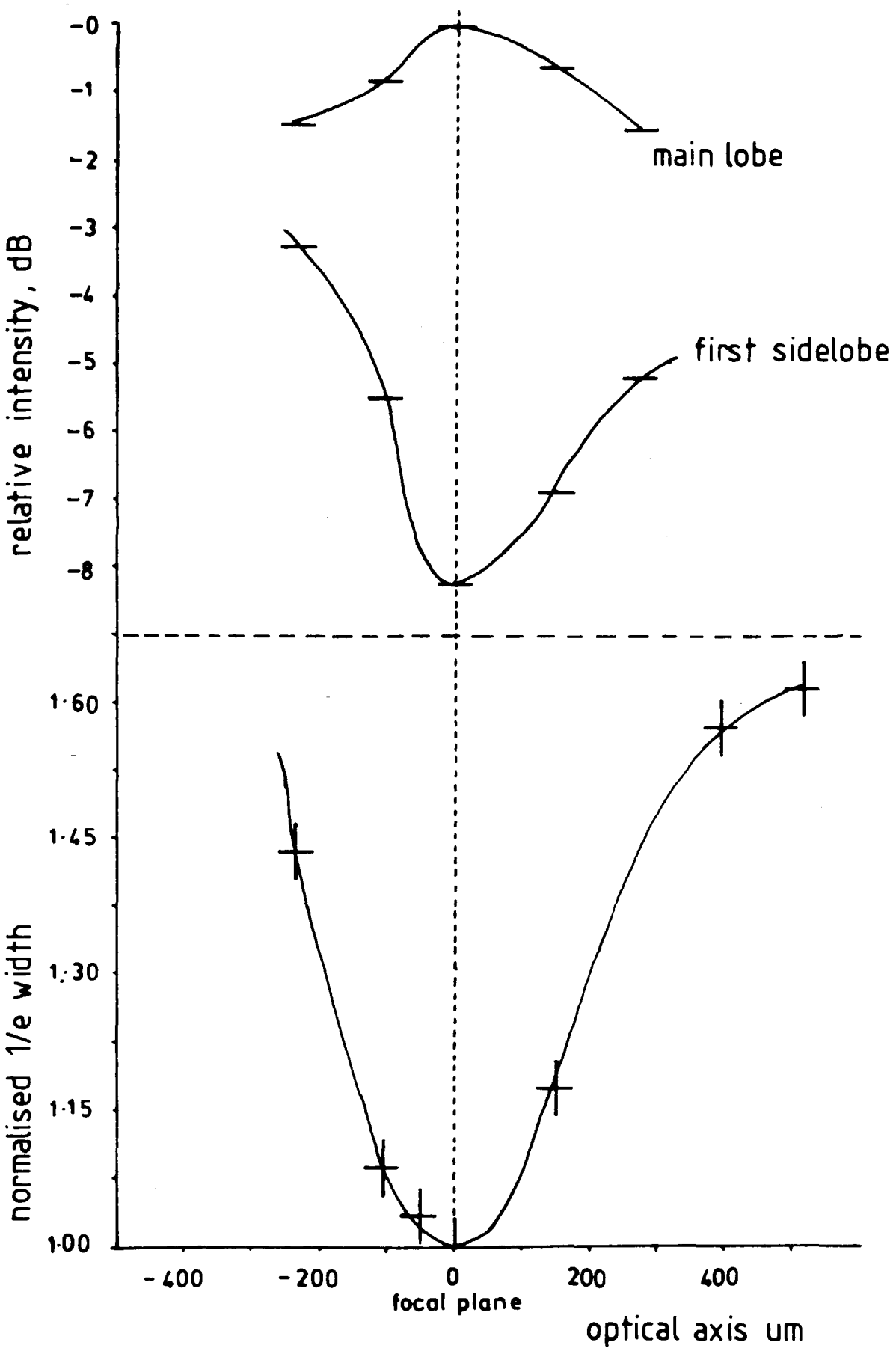


FIGURE [8,15]: Depth of focus measurements

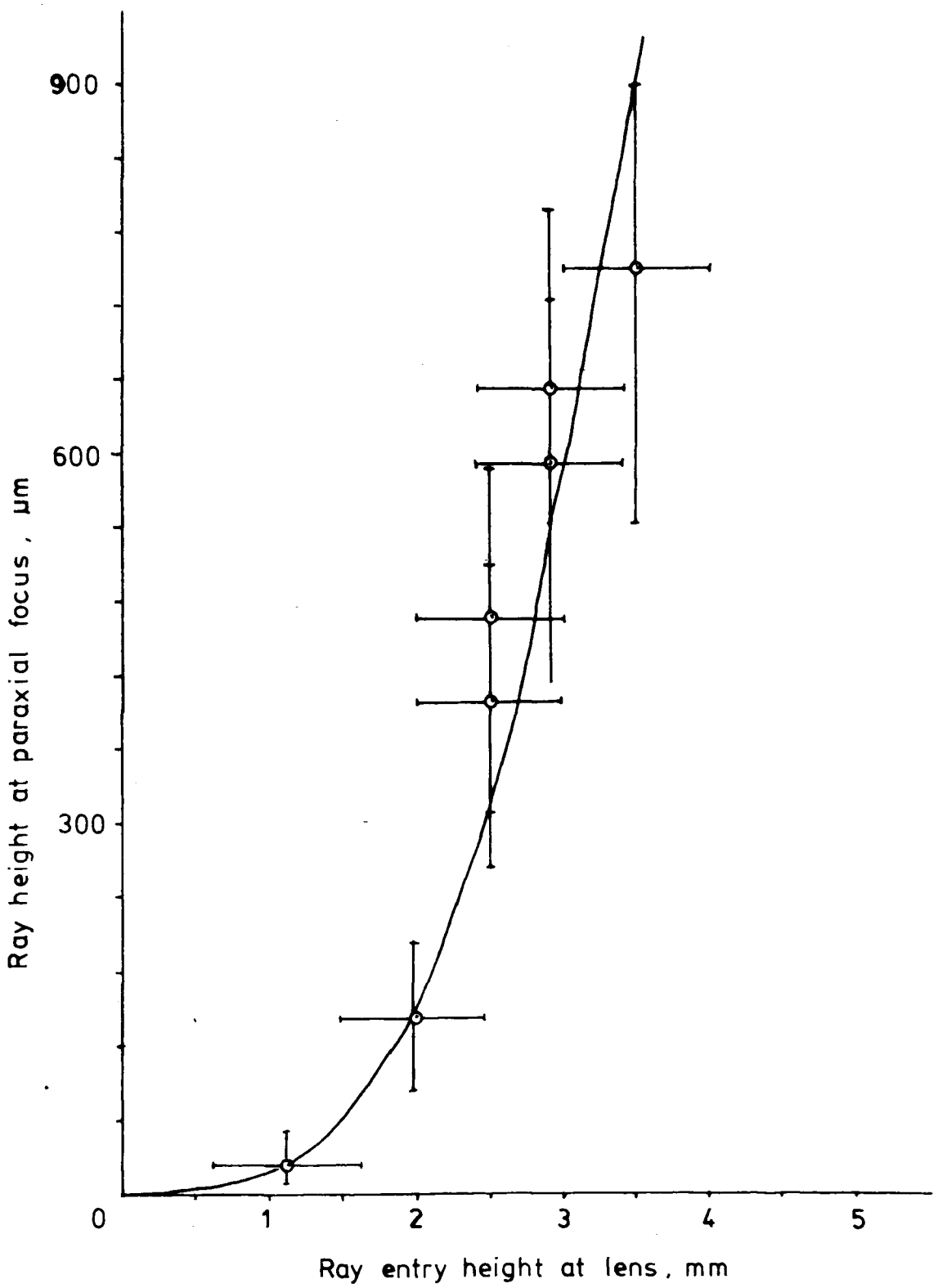


FIGURE [8,16] : Lateral component of spherical aberration

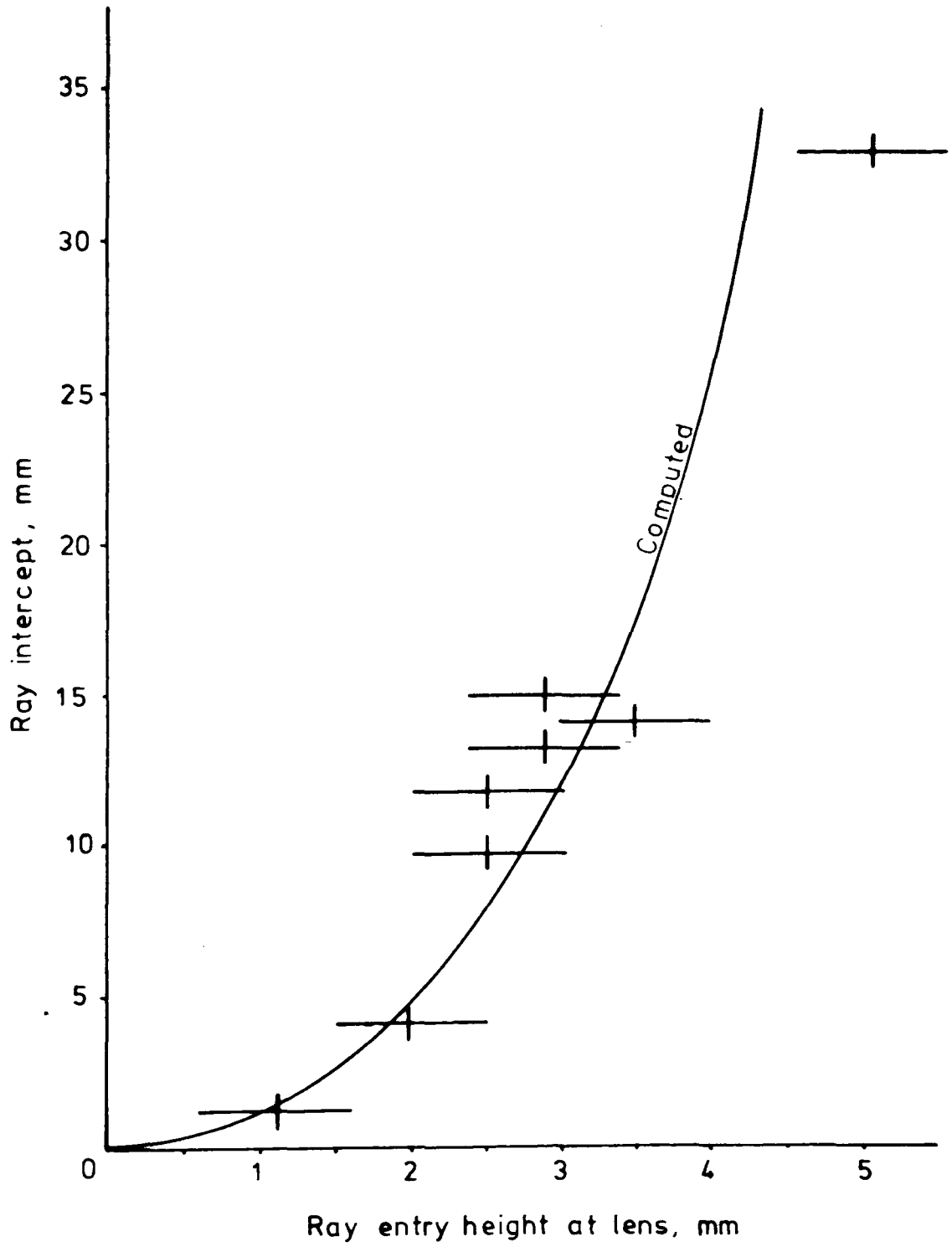
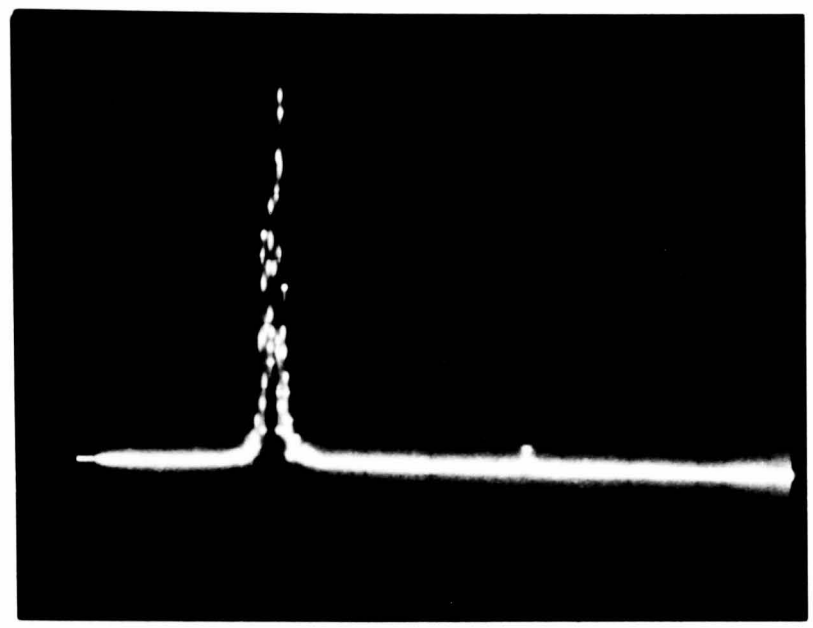
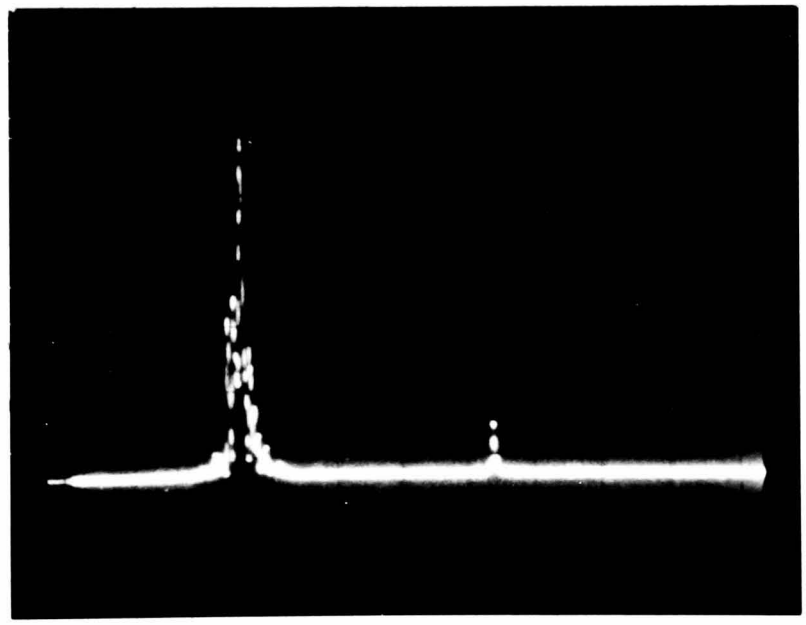


FIGURE [8,17] Longitudinal component of spherical aberration



(a)



(b)

FIG Reticon array images of undeflected and 1st order optical beams focused by geodesic lens after SAW diffraction [8,18]

(a) RF power = 56 mw (b) 141 mw

electrical power. The diffraction efficiency thus obtained was found to be 10-11.5%/watt.

The broad-band diffraction band-shape in the focal plane for two settings of input optical beam height are shown in Figure [8.19]. The differences may be accounted for, possibly, by the evolution of the acoustic beam from a Fresnel to a Fraunhofer regime, since "ringing" characteristic of Fresnel diffraction is observed in the upper photograph which appears to be absent in the lower photograph. A half-power bandwidth of 88MHz was observed.

For the narrow-band resolution test an untruncated, unexpanded beam with a $1/e^2$-diameter of approximately 2.0mm was used. Figure [8.20] shows an achieved resolution of 2MHz at a centre frequency of 746MHz. A re-imaging lens had been used to allow a 'Reticon' photodetector array to access the signals. The magnification factor of the re-imaging system was optimised to allow the separate diffracted signals to be detected on the array. Each element of the array has an 'active' region 15 μm wide, and adjoining elements are separated by a 'dead-band' 10 μm wide. Sparrow's criterion [18] was invoked in this experiment. The criterion states that if it is possible to detect the presence of two signals, without necessarily observing a dip between them (as Rayleigh's criterion requires), then the two signals are resolved.

8.7.8 Discussion on achieved resolution

The assumption is made that the re-imaging lens does not improve the angular resolution, since it would be expected that two diffracted signals which were not resolved in the front-focal plane of a lens would remain unresolved in the back-focal plane. Furthermore, the following calculations are for propagation wholly in the substrate, for convenience. While such was not the case, the numbers serve to illustrate the quantities involved.

The achieved resolution of 2MHz in the acousto-optic tests

corresponds to an achieved angular resolution in the substrate medium of 0.173mrad using the formula given in reference [6]:

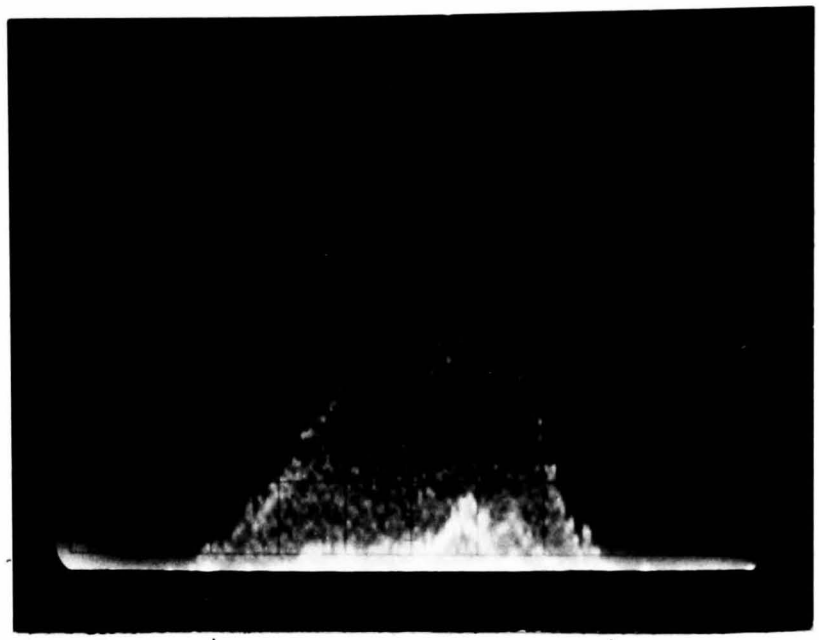
$$\delta\Theta = \frac{\lambda \delta f_s}{n v_r} \qquad (8.16)$$

where δf_s is the smallest resolved acoustic frequency, n is the effective refractive index encountered by the guided optical wave, λ is the free-space wavelength, and v_r is the velocity of the surface-acoustic wave. n = 2.21, v_r = 3488m/s and λ = 0.6328 μm were used to calculate the angular resolution. When multiplied by the focal length (f=50.9mm), the angular resolution corresponds to a separation between the spots arising from the two acoustic signals of 8.79 μm. Now the classical criterion for resolution, given by Lord Rayleigh [8], is:

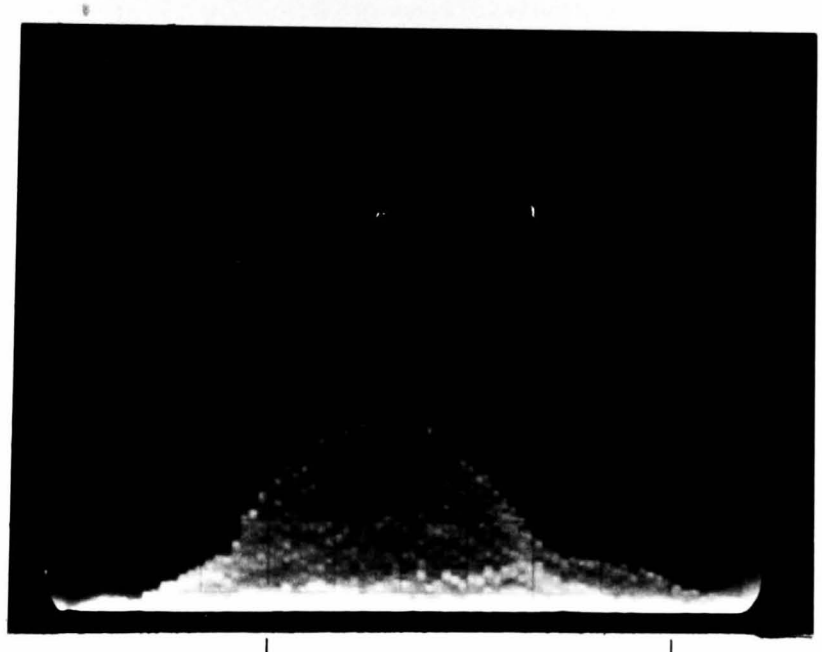
$$\text{Separation} = \frac{0.82\,\lambda}{NA} \qquad (8.17)$$

for coherent illumination, where NA = nD/2f is the approximate value of the numerical aperture calculated using the effective beam-diameter D, which in the experiment was 2.0mm. Inserting the appropriate values in (8.17) gives a smallest separable interval of 12.6μm. Thus a resolution better than the Rayleigh limit was apparently achieved. In reality, the Rayleigh 'limit' is not a true limit at all, and the scale factor of 0.82 is somewhat arbitrary, depending on the actual form of the input optical signal. The signals were not resolved according to Rayleigh's criterion since no dip was observed between them. Sparrow's criterion may be invoked instead. The reason for this choice is that adjacent, discrete detector elements in an array were used in the experiment to pick up the diffracted optical signals. Clearly, no 'dip' could be observed in such a case. The minimum resolvable interval according to the Sparrow criterion is:

$$\text{Separation} = \frac{0.5\lambda}{NA} \qquad (8.18)$$



(a)

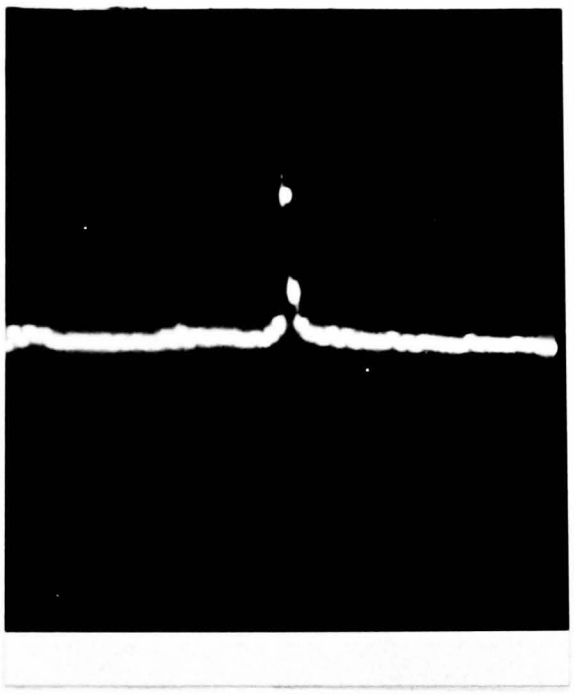


(b)

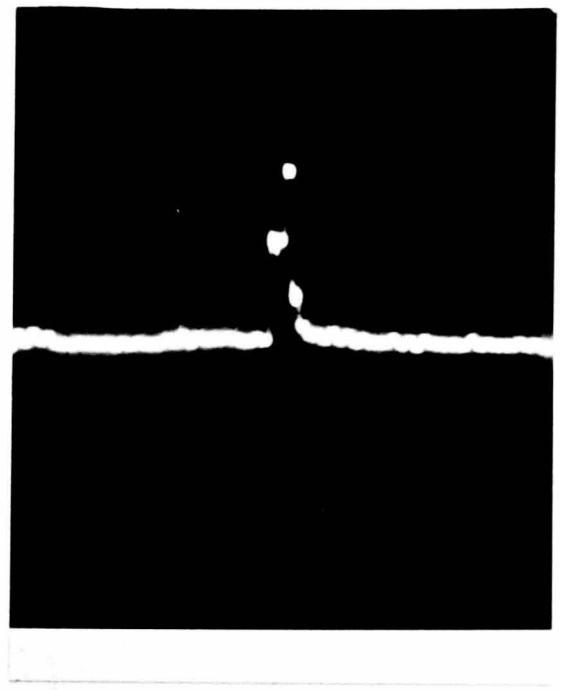
FIG A/O diffraction bandshape with
[8,19]

(a) 1mm optical beam, traversing centre of lens

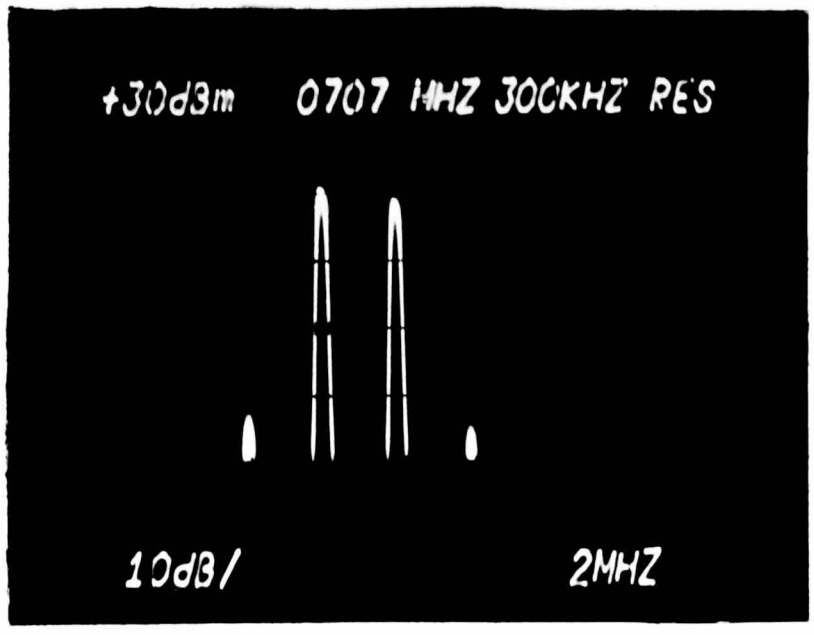
(b) beam 1mm further away from transducer



(a)



(b)



(c)

FIG [8,20] Resolution of signals 2MHz apart by IOSA

(a) single signal

(b) 2nd signal introduced, 1st signal unaffected

(c) electrical spectrum of the two signals

which is 7.29μm in this, the present case. The experimental resolution was, therefore, only just above the theoretical minimum, indicating that near-diffraction limited performance is possible using spherical geodesic lenses, albeit at considerably restricted relative apertures.

The above result was obtained at optimum acoustic frequencies. At other frequencies the resolution of the system decreased considerably, so that the diffracted signals were spread over several detector pixels at the same system magnification as used above. Proper broad-band characterisation of the device has not yet been carried out, although the device is equipped with SAW transducers which should allow broad-band operation.

8.8 Overlay lens LL1

8.8.1 Results

A photograph showing two pencils of light entering lens LL1, leaving the lens and intersecting at an axial point was obtained. The light levels in the experiment were very low except at the entrance and exit surfaces of the lens, where a great deal of light was scattered out of the plane of the waveguide due, probably, to the fairly abrupt interface that existed between the overlayer and the ambient waveguide.

Lens LL1 was 4.25mm in diameter. The two light pencils in the photograph were separated by 2.0mm at the entrance to the lens, and each pencil had a diameter of 0.25mm. The intersection point of the two pencils along the optical axis was 6.4mm from the centre of the lens. A bright, diffuse glow was evident in the photograph at this point. The lens operated at an f-number of f/3.2 . A true Luneburg lens having the same diameter and thickness at the centre as the fabricated lens would have

focused at 8.5mm. The lens, therefore, introduced a negative longitudinal component of spherical aberration of 2.1mm .

Light pencils spaced more than 2.5mm apart did not propagate through the lens to the focus, probably due to the small angles of incidence with the lens at the margins. When a beam of light was used, rather than two narrow pencils, the focal region became a very diffuse patch, extending to several tenths of mm in the transverse direction and between one and two mm in the direction of propagation. This observation suggested that the paraxial portion of the beam was focusing at further distances than the marginal portions. The spotsizes produced by this lens and other ovarlay lenses were so large that they were not considered worth measuring in detail.

8.8.2 Discussion

The observed behaviour of lens LL1 was very similar to that predicted by the ray-tracing analysis presented in chapter four for the 'best-obtainable' approximation to the true Luneburg lens. The actual fabricated profile differed from the best approximation and the true Luneburg lens in that it was thicker than both over most of the profile. The fabricated lens would be expected to be under-corrected with respect to the true lens, in similar fashion to the best approximation, only more so. Such was the case. The focal length of the best approximation was predicted to be smaller than that of the true Luneburg lens in the ratio 3.2/4.0, or 0.8 . In the case just examined, the ratio of the measured focal length to the desired focal length was 6.4/8.5 or 0.75 . Furthermore, the paraxial rays also behaved as predicted in chapter four.

The results, whilst expected, were disappointing, for As_2S_3 is one of the few high-index materials available for use with $LiNbO_3$.

8.9 Geodesic lens GL1

8.9.1 Photographs of irradiance patterns

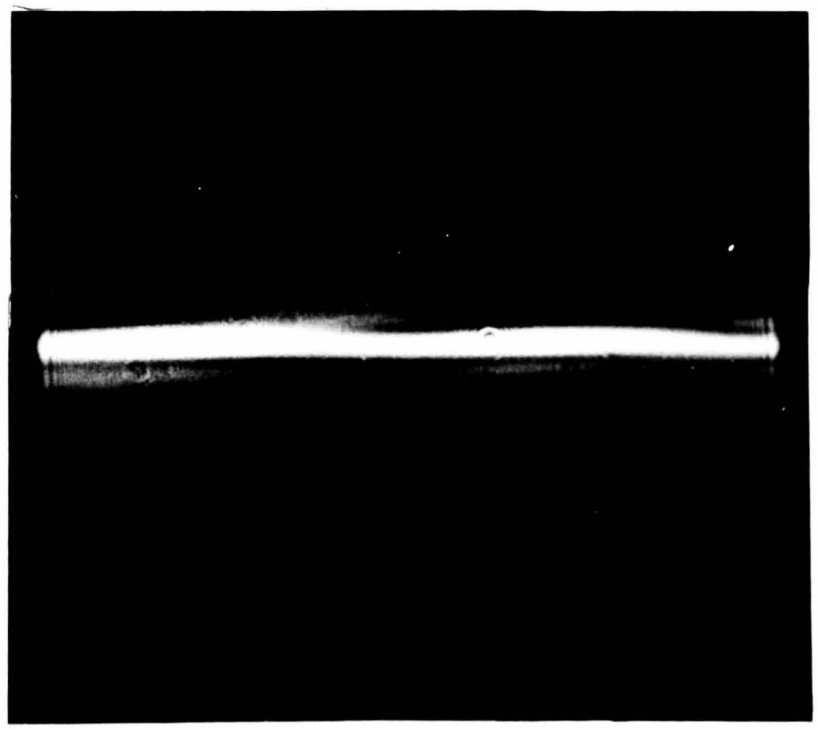
A photograph of the focal-plane image irradiance of lens GL1, a theoretically perfect geodesic lens of aspheric design, is shown under conditions of under-development, normal-development and over-development in Figure [8.22] (b), (a) and Figure [8.23] respectively. The photograph was obtained by G.F. Doughty and is represented as Figure [7.17] on page 222 of reference [19]. Showing the irradiance pattern at the focus under different conditions of development of the photograph serves to highlight different aspects of the image, from the well-represented central lobe in the under-developed case, to well-represented sidelobe structure in the over-developed case.

An expanded Gaussian beam having a $1/e^2$-diameter of 4.0mm was truncated at a diameter of 3.0mm and used as the incident beam on the lens. The beam was coupled into and out of the waveguide using rutile prisms. The sidelobes arising as a result of truncation were very prominent, as seen in Figure [8.23]. Using equation (8.14), the width between the nulls of the image in the diffraction-limited case would have been 3.54 μm. The experimental value, measured with a calibrated microscope, was 5.0 μm $\pm$ 0.5 μm.

8.9.2 A computer-acquired image

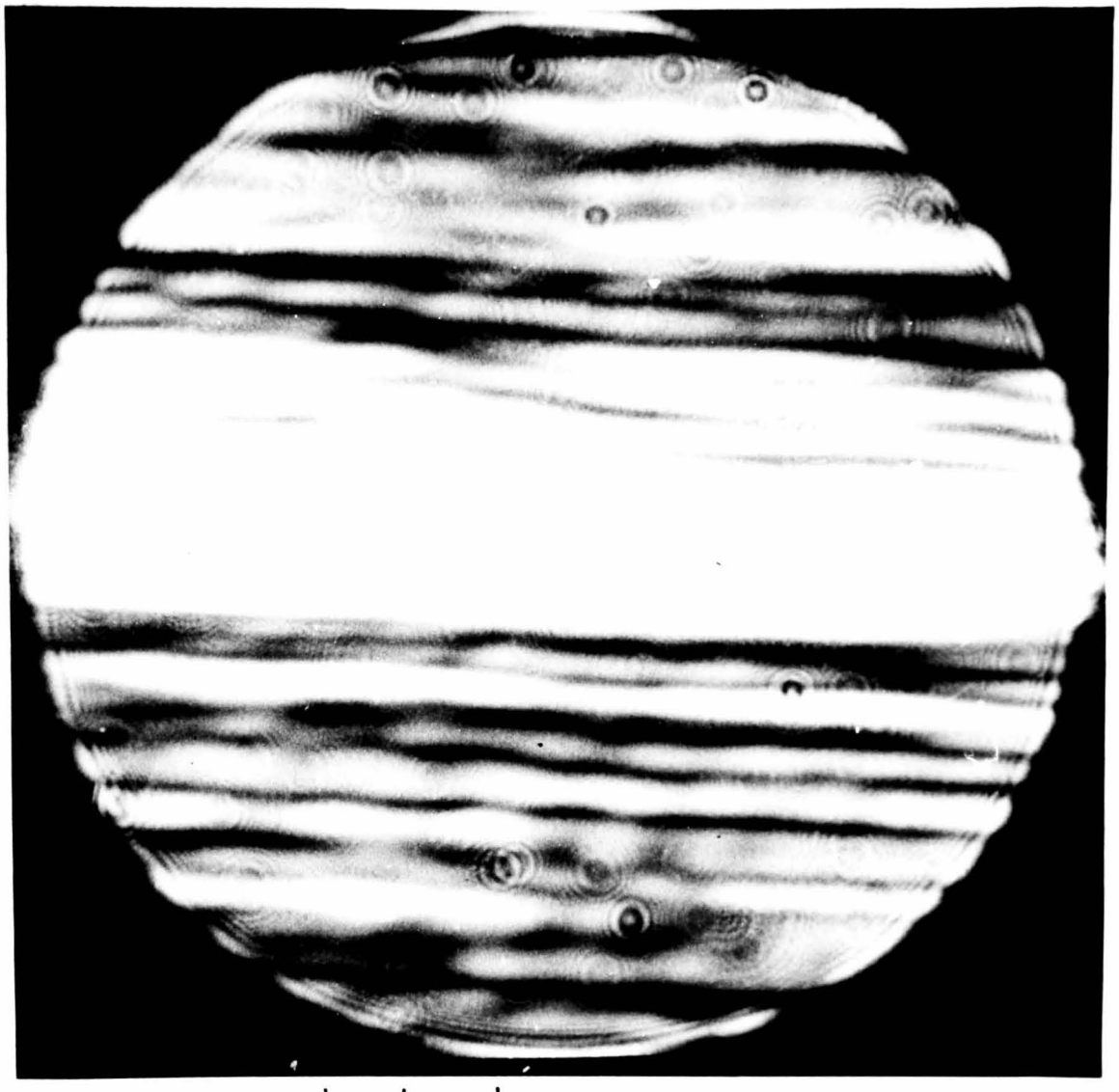
A Gaussian beam having a $1/e^2$-diameter of 2.1 mm was coupled to the waveguide and lens GL1 using a rutile prism. The beam had not been truncated outside the waveguide. Instead of coupling the light out of the waveguide with a prism after it had passed through GL1, the beam was allowed to propagate to the polished edge at the end of the substrate. The irradiance profile at the edge was then projected onto a Newvicon vidicon and recorded by the Hammamatsu image-acquisition system. A contour map of the image is shown in Figure [8.24] (a) . A bright central spot is

(a) developed



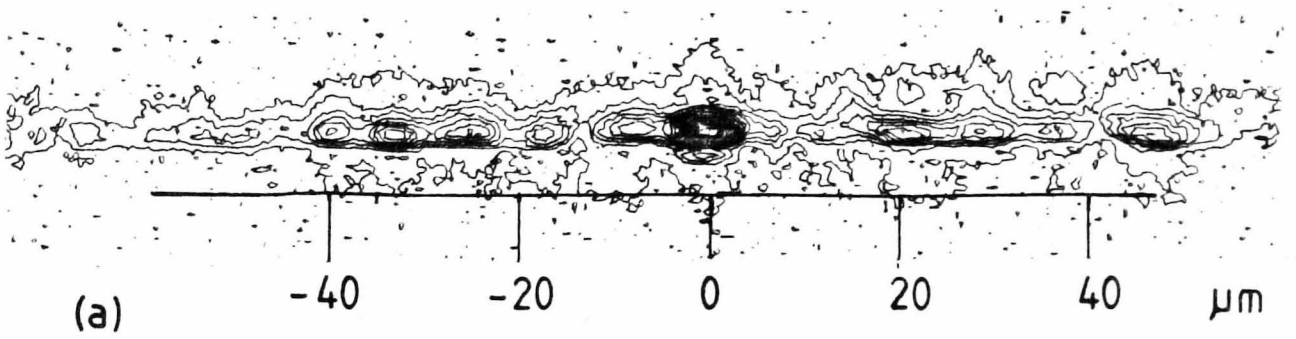
(b) under-developed

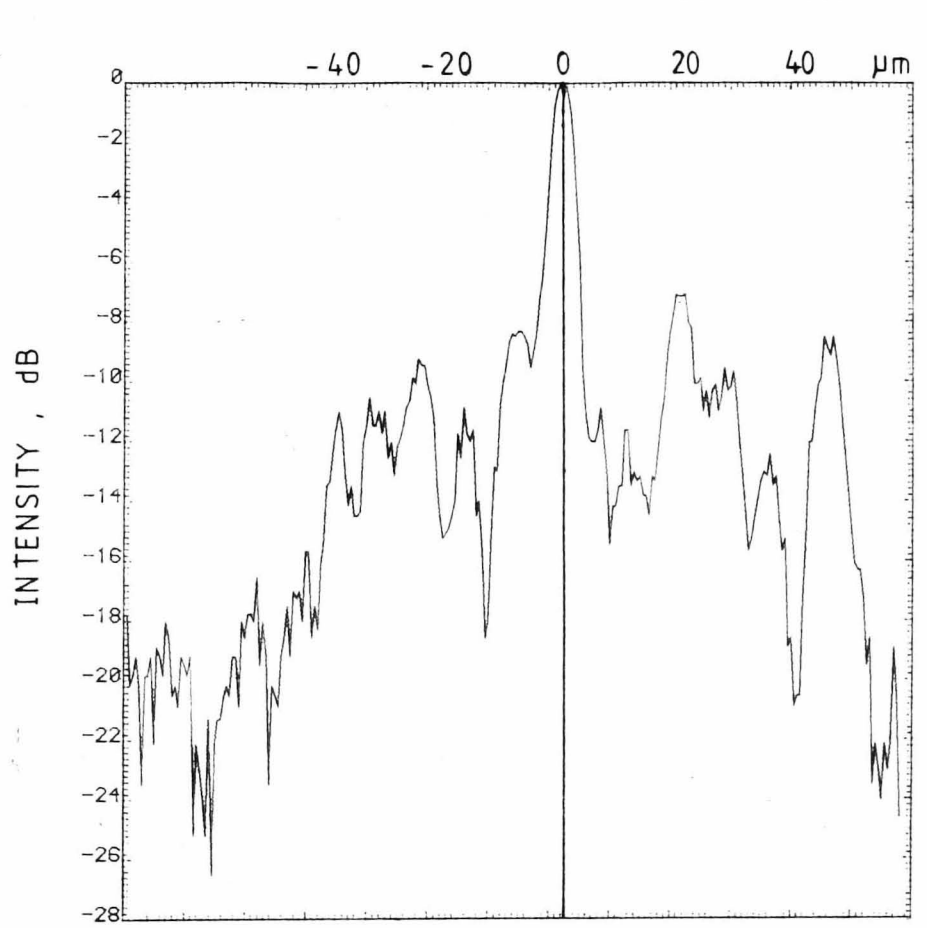
FIGURE [8, 22] Focal plane irradiance patterns of aspheric geodesic lens



over - developed

Figure [8,23] Focal plane of Sottini geodesic lens





(b)

FIGURE [8,24] (a) GL1: ISOPHOTES AT FOCUS

(b) IRRADIANCE AT FOCUS

evident, together with a considerable amount of sidelobe structure. The presence of sidelobes indicated that the input beam had not remained Gaussian, although the width of the beam at the input was considerably smaller than the limiting apertures of the input prism and the lens itself.

A line-scan of the irradiance across the range of peaks of Figure [8.24] (a) is shown in part (b) of the same Figure. The spatial scale is slightly reduced. The $1/e^2$-diameterof the central lobe is 7.9 um and the width between the nulls is approximately 11.0 μm. Using equations (8.14) and (8.15), the corresponding theoretical widths for a uniformly-illuminated aperture are 3.54 μm and 5.1 μm respectively. The sidelobes are seen to be very high, with one sidelobe only 7 dB below the peak value. The sidelobes are also asymmetric and exhibit an irregular decay. Clearly, the lens introduced amplitude- and/or phase-distortions into the propagating wave-field. These could be attributed to in-plane scattering, position-dependent curvature and leaky-mode loss or non-uniformity of the waveguide over the lens region.

8.9.3 Direct measurements of the irradiance profiles

The experimental set-up shown in Figure [8.6] was used to obtain further measurements at a wider range of beam apertures. An expanded Gaussian beam having a $1/e^2$-diameter of 7.2mm was used. Irradiance profiles measured at the edge of the substrate for truncated beamwidths of 0.47 mm, 1.02 mm, 1.86 mm, 2.78 mm, 3.16 mm, 3.22 mm, 4.18 mm and 6.97 mm are shown in Figure [8.25], parts (a) to (h) respectively. Also plotted on these graphs are sinc^2 curves having the same $1/e^2$-diameter as the experimental curves. The sinc^2 curves are, therefore, not the diffraction-limited curves. The sinc^2 curves are presented solely to highlight aspects of the sidelobe structure. A reasonable justification for this procedure is that the aperture external to the waveguide may not be the limiting aperture. If either the coupling prism or the lens itself has a transmission function with a width effectively smaller than the limiting aperture, the

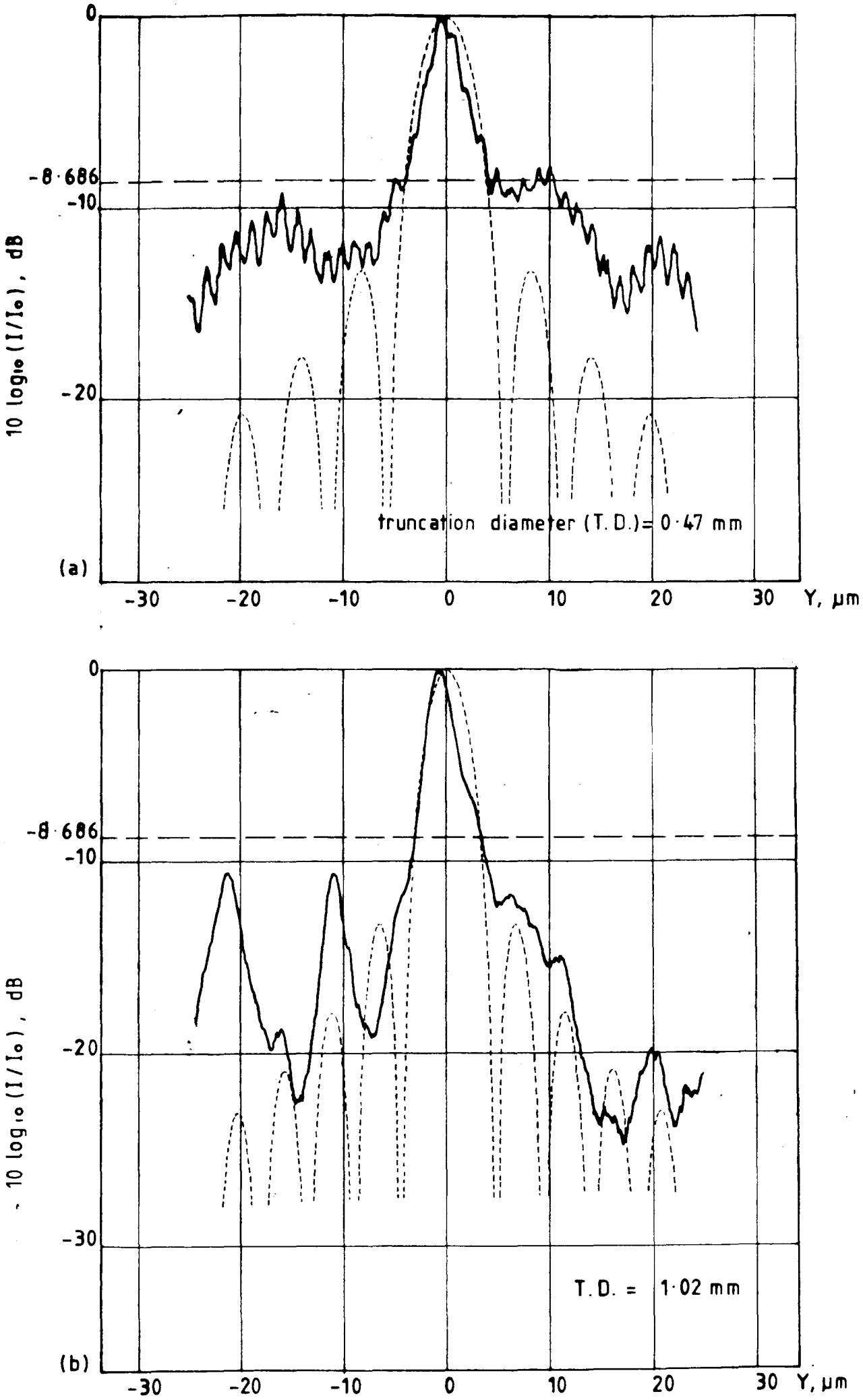


FIGURE [8,25] Irradiance profiles at focal plane of lens GL1 at various values of aperture truncation. $(Sin(y)/y)^2$ curves having same $1/e^2$ widths as experimental profiles are also shown.

(a) truncation width = 0·47 mm
(b) " " = 1·02 mm

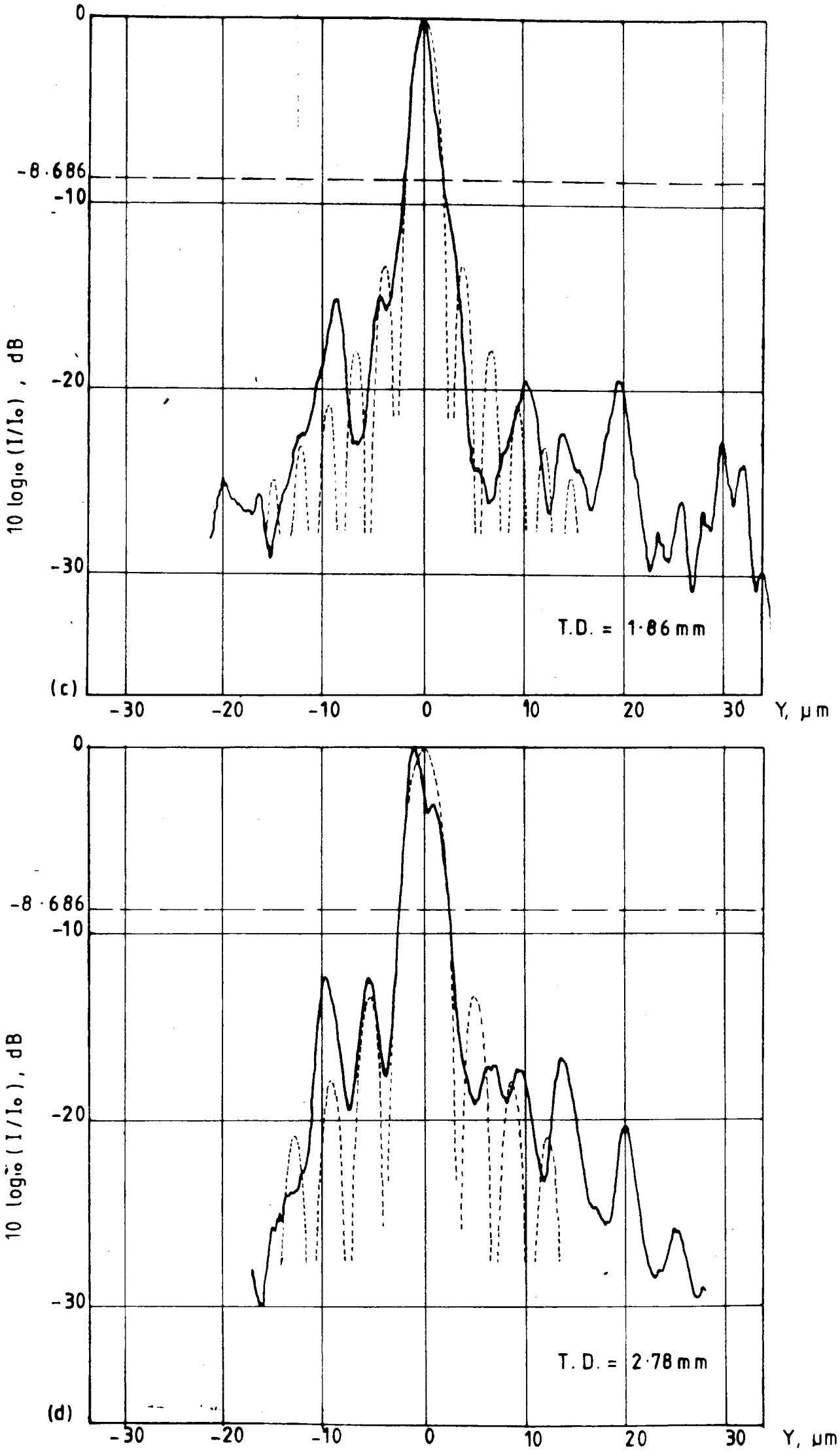


FIGURE [8,25]
(c) truncation width = 1·86 mm
(d) " " = 2·78 mm

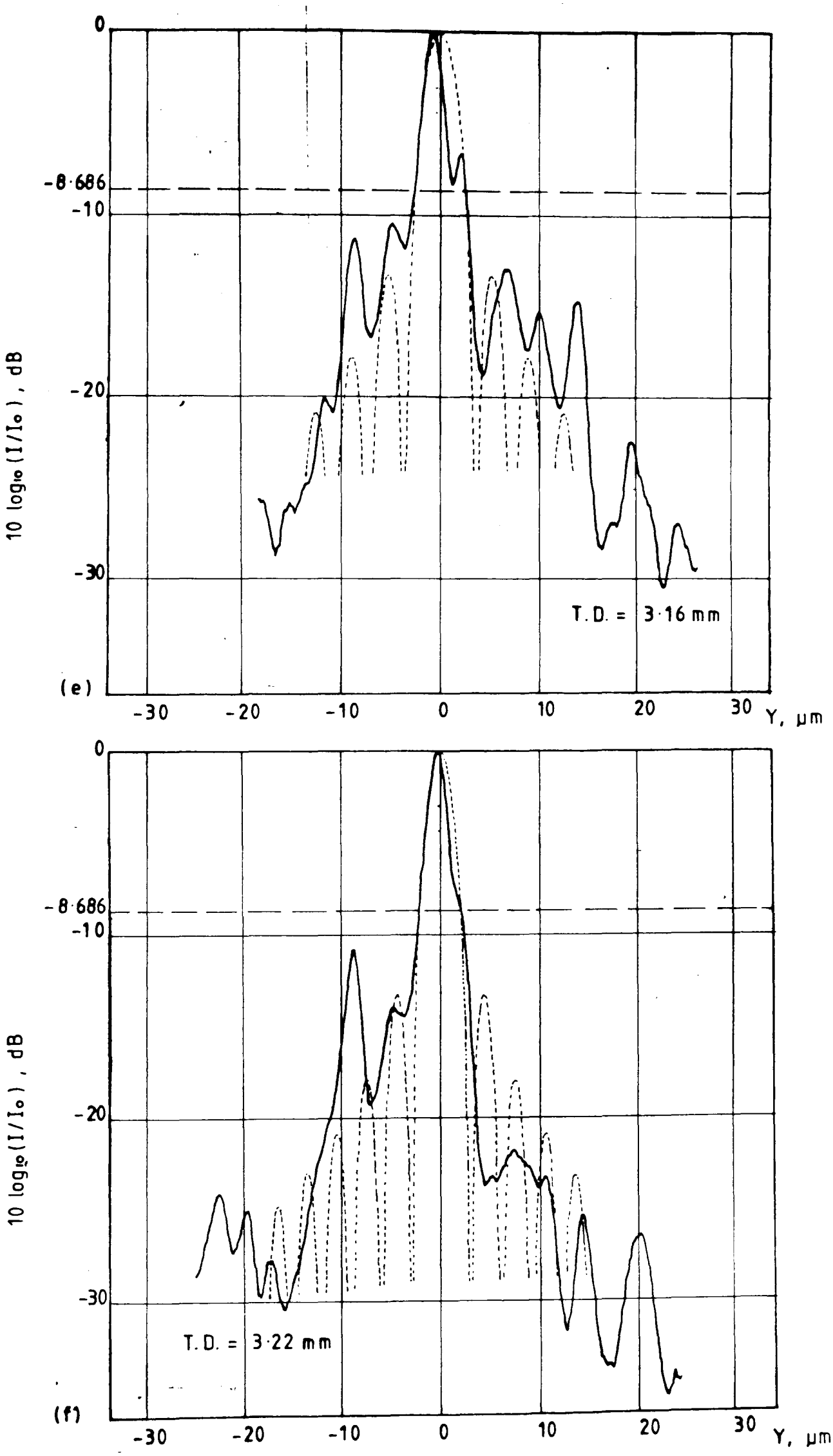


FIGURE [8,25] (e) truncation width = 3·16 mm
(f) truncation width = 3·22 mm

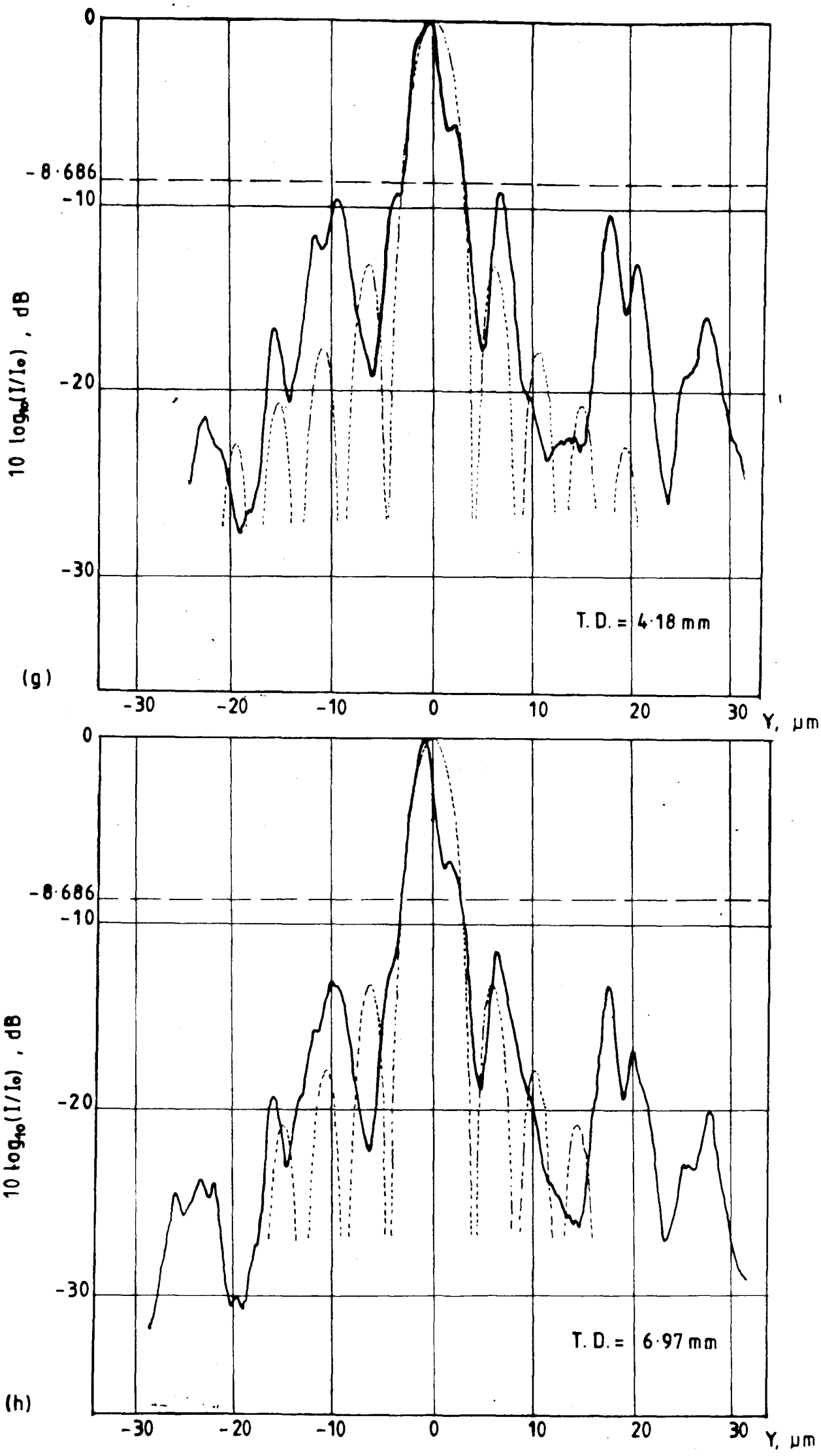


FIGURE [8,25] (g) truncation width = 4·18 mm
(h) " " = 6·97 mm

diffraction pattern at the focus would be broadened accordingly. The coupling prism, in particular, is unlikely to have a step transmission function.

Some observations may be made about the measured irradiance profiles:

(i) they are all asymmetrical with respect to the central spot, with the sidelobes on the left of the patterns, as drawn, being higher than those on the right, in general. Such asymmetry may arise from:

(a) local scattering defects in the vicinity of the focal region, either in the planar waveguide, or at the end-face;

(b) a deviation of the input beam from axial incidence, in which case the asymmetry could indicate the presence of coma;

(c) an asymmetrical titanium evaporation at the waveguide formation stage.

(ii) The sidelobe structure is not well-ordered in that a regular decay is not exhibited, indicating that phase-distortions were imposed on the optical field;

(iii) the sidelobe structure varies quite dramatically with aperture width, as illustrated in the following Table. The variation of sidelobe height as a function of aperture width appears to be fairly random and asymmetric.

Aperture Width,mm	Left Sidelobe,dB	Right Sidelobe,dB
0.47	-9.2	-7.9
1.02	-10.5	-11.7
1.86	-14.9	-19.4
2.78	-12.4	-16.8
3.16	-10.5	-6.7
3.22	-14.0	-21.6
4.18	-9.2	-5.7
6.97	-13.0	-6.7

TABLE [8.2] Height of first sidelobes in Figure [8.25] (a)-(h)

(iv) At an input aperture of 1.86 mm, the lens exhibits diffraction-limited performance, both in terms of central spotsize, and adjoining sidelobe height. The effective f-number at this aperture is approximately f/10.

The variation of the half-power spot-diameter, the 1/e-diameter and the $1/e^2$-diameter is shown in Table [8.3] and plotted in Figure [8.26]. At f/10, these spot-sizes are 1.9 μm, 2.9 μm and 3.8 μm respectively. At an input beamwidth of 3.22 mm, corresponding to f/5.75, the $1/e^2$-diameter is 4.2 μm, twice the diffraction-limited value. As with the sidelobe structure, the variation in spotsize with input aperture width follows an irregular development, as seen in Figure [8.27], where the $1/e^2$-diameter results, a parabolic least-squares fit and the diffraction-limited curve are plotted together. Clearly, a parabolic fit to the experimental data is not a good representation of spotsize behaviour.

8.10 Substrate GL2

The substrate GL2 supported two geodesic lenses intended to approximate closely the perfect profiles. As was shown in Figure [8.2], the end-polishing process had reduced the distances between the centres of the lenses and the ends of the substrates to 18.47 mm and 18.32 mm for lenses designated GL2(a) and GL2(b)

TABLE [8,3] Experimentally-observed focal spot-sizes as a function of size of (truncated) input be for lens GL1

Truncated beam-diameter, mm	Focal spot-diameter, μm			Diffraction-limit, μm (1)
	-3 dB	1/e	$1/e^2$	
0.47	3.4	4.7	8.1	14.4
1.02	2.8	3.4	6.4	6.6
1.86	1.9	2.9	3.8	3.6
2.78	1.7	3.6	5.1	2.4
3.16	1.9	2.4	5.1	2.1
3.22	1.9	2.3	4.2	2.1
4.18	2.9	3.3	6.2	1.6
6.97	2.1	2.8	6.0	1.0

NOTES: (1) diffraction-limit is calculated for a Gaussian beam

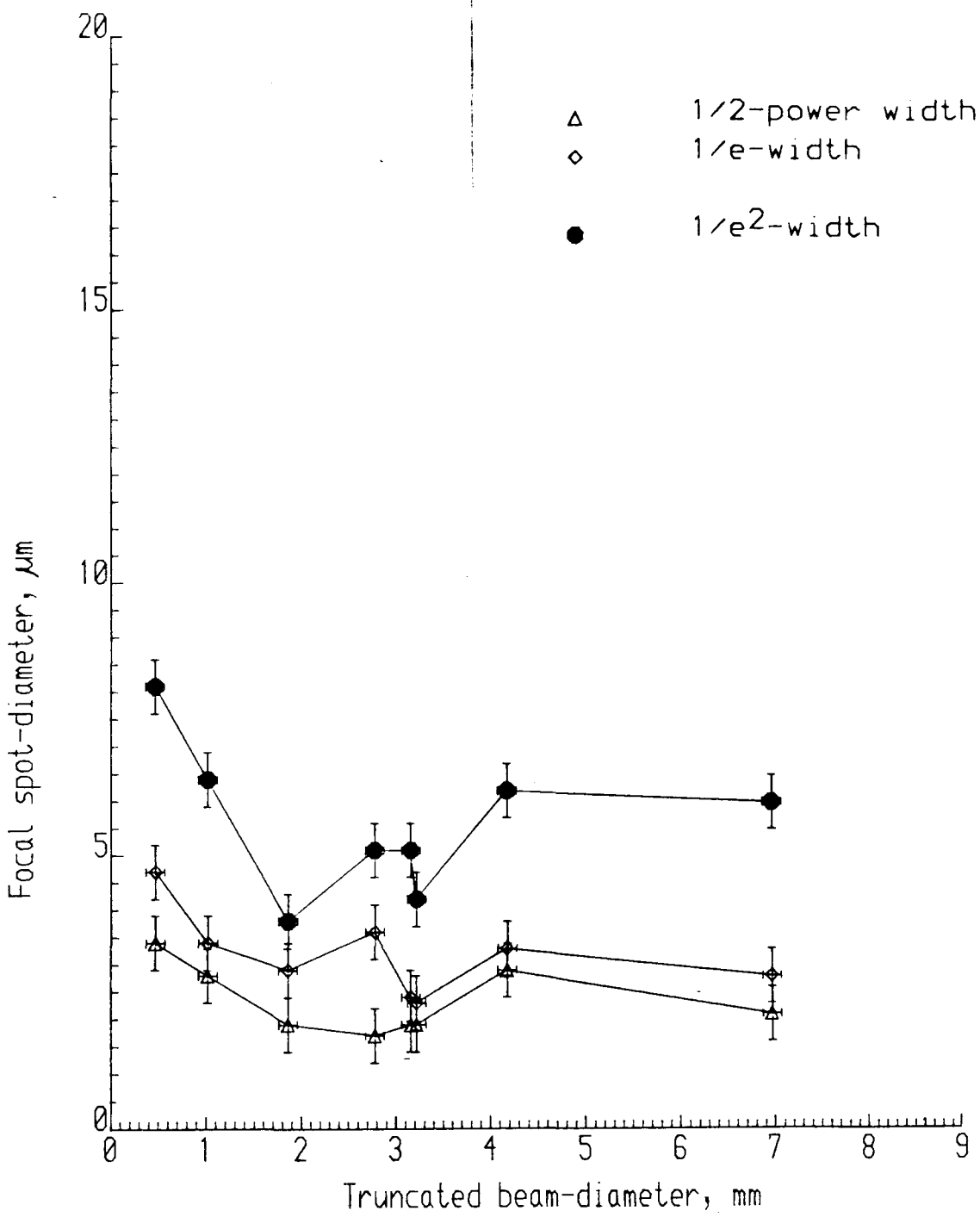


FIGURE **[8,26]** Measured sizes

of image-plane focal spots

of lens GL1

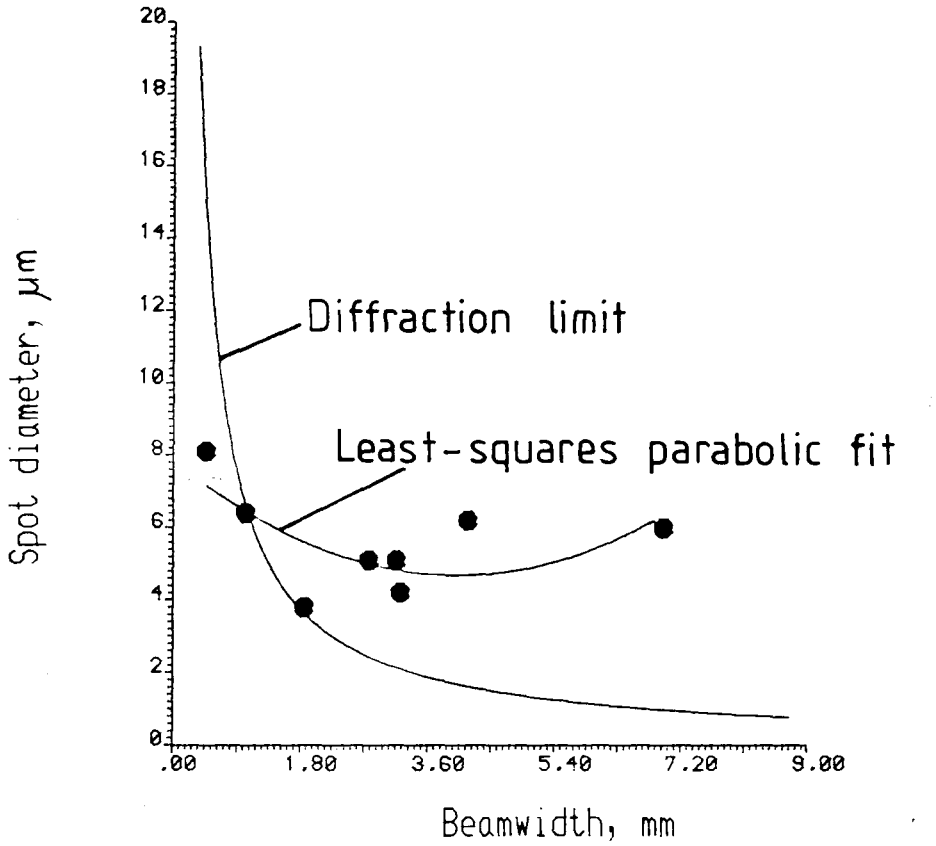


FIGURE [8,27] GL1 Focal spotsizes

respectively. The design focal length was 18.5 mm. The substrate is shown under test in Figure [8.28] .

A substrate of this type, supporting two geodesic lenses, was investigated in chapter seven of this thesis in an investigation into the uniformity of depth likely to be obtained in the waveguide . One of the conclusions of that chapter was that under certain circumstances, the waveguide in the lens regions could fall below cut-off if the waveguide in the planar region was not far into the single-mode regime. The conclusion was confirmed in studying GL2. A layer of titanium, 230 A thick at the centre of the substrate, was deposited from a distance of 220 mm. The metal was diffused into the substrate for 10 hours in a wet argon atmosphere at 1000^{o}C, followed by cooling in a wet oxygen atmosphere. The waveguide formed on the planar region of the substrate supported one TE mode only, with an effective refractive index of 2.206 measured using prism-coupling techniques. The lenses were tested and found to transmit over most of their apertures, except for 1 mm bands in the centres. A further layer of titanium, 100 A thick at the centre of the substrate, was consequently deposited and diffused into the substrate. The waveguide in the planar region was then found to support two TE modes, with effective indices of 2.212 and 2.207. The lenses were now found to transmit guided light for the lowest-order mode, though the other mode failed to propagate. The shape of the transmission function of the lenses was not measured. Such a measurement would be useful and would best be carried out by sawing and polishing an end face close to a lens, to prevent focusing, planar waveguide propagation loss and diffraction from contributing to any irradiance changes. Ideally, the lens would be uniformly illuminated at the input (the entrance pupil). However, neither prism-coupling nor end-fire coupling is capable of doing this.

An irradiance profile, measured at the best focal distance, as estimated by the eye through a microscope, is shown in Figure [8.29]. The incident field was an untruncated gaussian with a $1/e^2$ beam-diameter of 3.74 mm. The focus occurred at 18.75 mm,

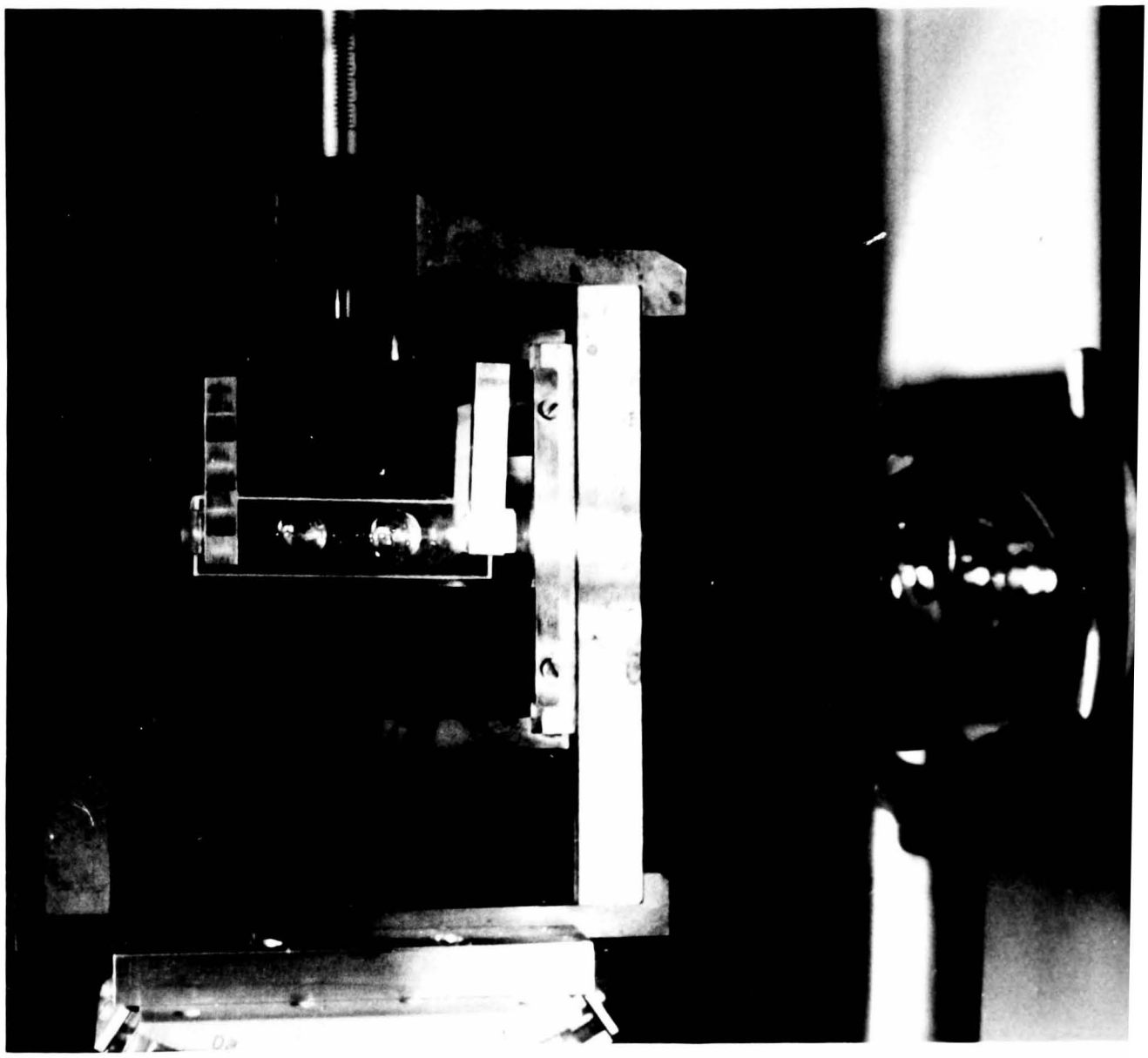


Figure [8,28] Two-lens system mounted for optical testing

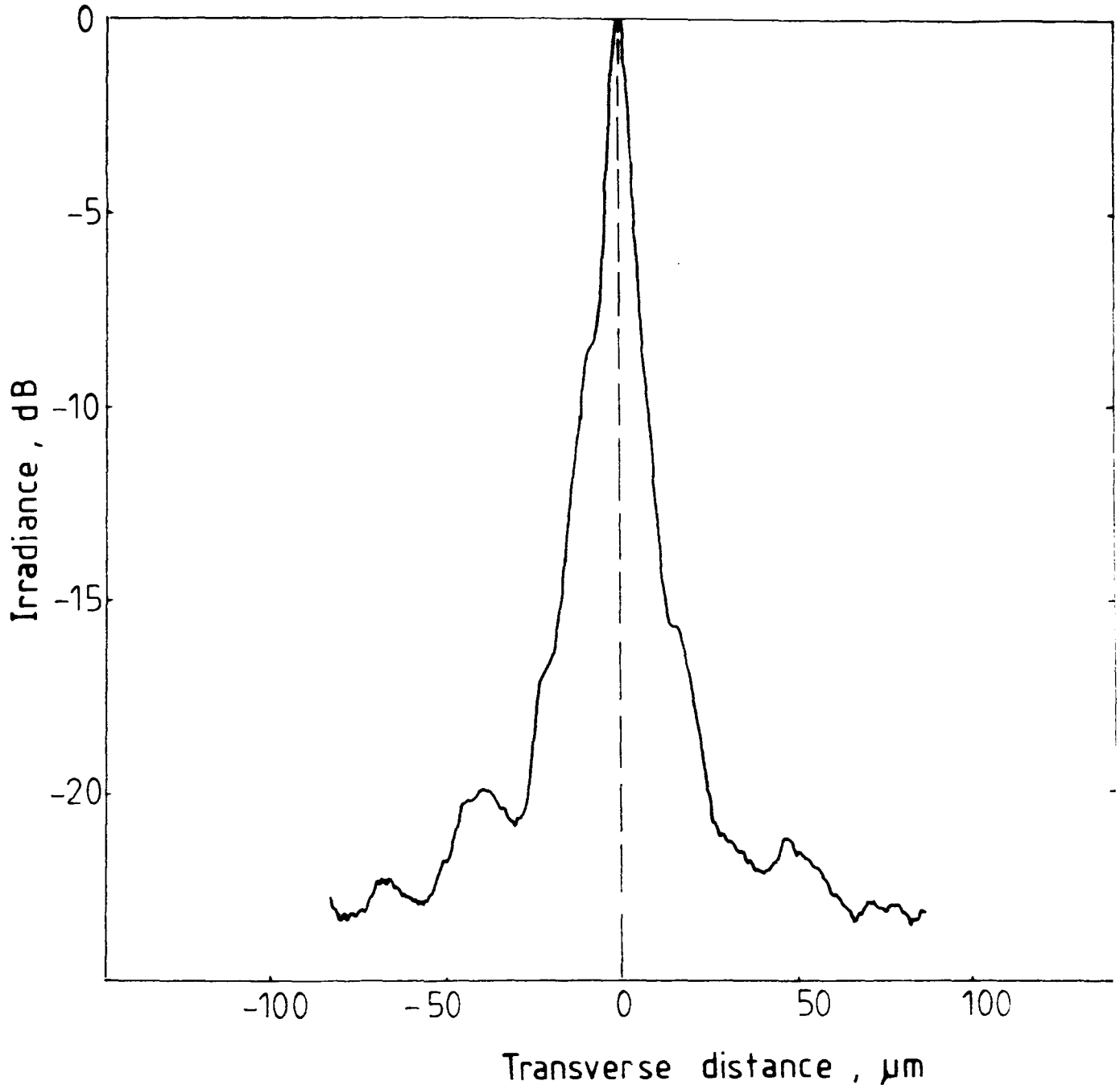


FIGURE [8,29] FOCAL PLANE IRRADIANCE PROFILE, GL

although focusing in air should have reduced the focal length below 18.5 mm. This possibly indicates that an objective assessment of the best focus is not possible with the human eye and that measurements of the irradiance patterns over the focal volume are always required. At the time the irradiance profile was obtained, the substrate dimensions had not been taken, and so focal volume irradiance profiles were not obtained.

The half-power width of the measured field was 6.1 um and the $1/e^2$-diameter was 14.3 μm. The diffraction limit for the lens for an incident field having a diameter of 3.74 mm is 1.8 μm, so that the measured irradiance profile was not diffraction-limited, possibly owing to a measurement at a distance beyond the true focus. It is apparent from the Figure that the field is quite smoothly shaped, displaying little scattering noise and sidelobes at less than 20 dB below the central peak.

The other lens on the substrate, GL2(b), displayed considerably degraded optical performance, as shown in Figure [8.30]. The irradiance profile at the estimated best-focus is shown. The incident field was a Gaussian having a $1/e^2$-diameter of 7.2 mm, truncated at 3.17 mm. The half-power width of the field shown in the Figure is less than 5 μm wide. The $1/e^2$-diameter is 17 μm, however, and a great deal of scatter is in evidence.

8.11 Focal length of lens SL1

In chapter five, the beam propagation method (BPM) was employed to calculate the optical field in the vicinity of the focus of diffraction-limited geodesic lenses. An important result of that chapter was that a shift of the focus occurred from the focal length predicted by geometrical optics. The shift occurred in a negative direction, ie towards the lens. This shift was also observed with a spherical geodesic lens having a toroidal edge-rounding region, a lens expected to produce positive spherical aberration. In chapter five it was speculated that the negative defocus that is observed using a wave-optics analysis such as the

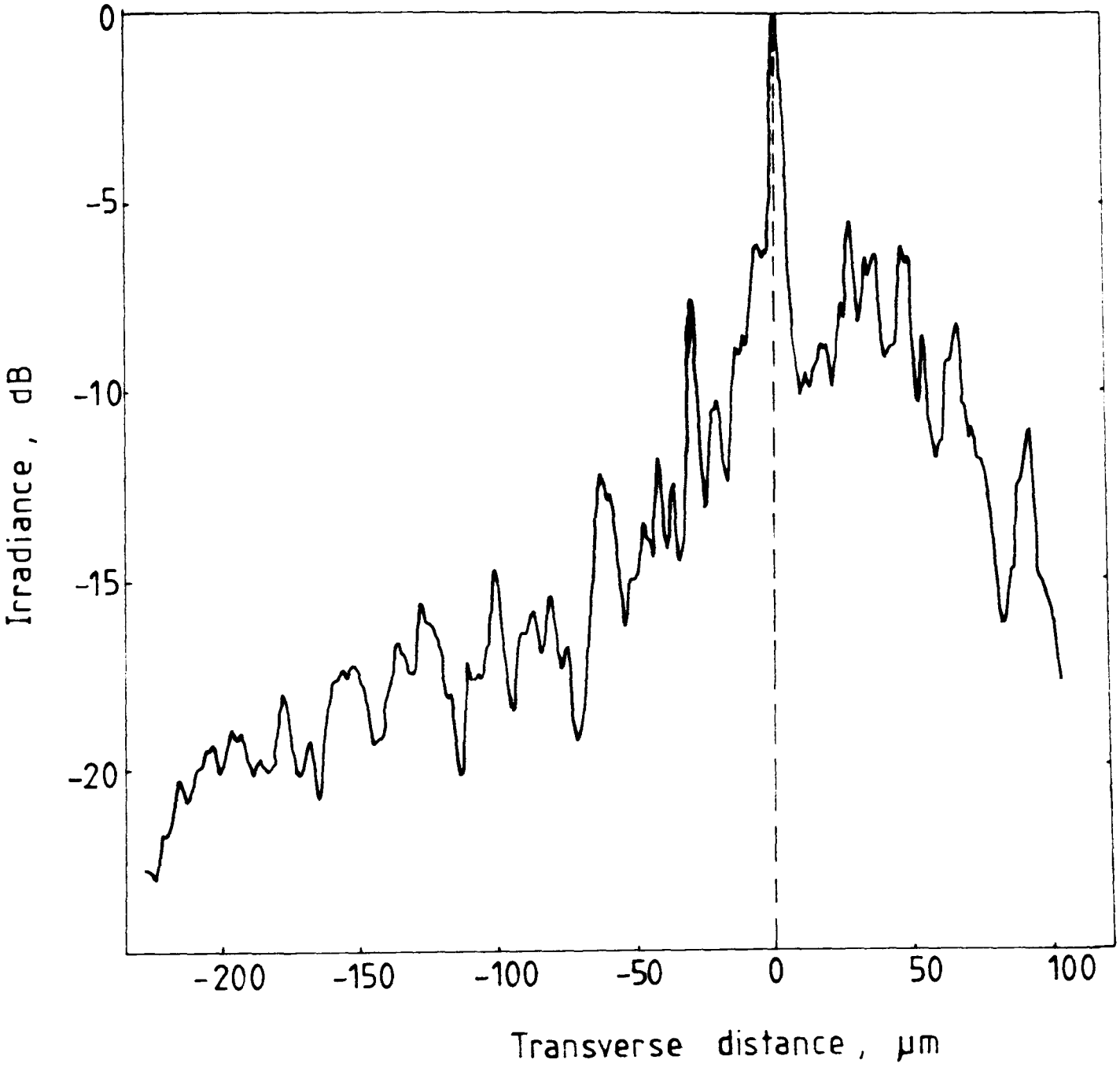


FIGURE [8,30] FOCAL PLANE IRRADIANCE PROFILE, GL2(b)

BPM would compensate, to a degree, for the spherical aberrations predicted by the purely geometrical theory.

Experimental evidence appears to confirm that such aberration compensation occurs. Figure [8.31] shows measured estimates of the focal length of lens SL1 as a function of aperture width. The field incident on the variable aperture was approximately Gaussian, with a $1/e^2$ diameter of 10.0 mm . Two geometrical curves are also shown in Figure [8.31], calculated from ray-tracing. One geometrical curve was calculated with a ray manifold that had a Gaussian weighting and truncation imposed upon it, to simulate experimental conditions. The other geometrical curve was calculated with a uniform weighting. In both cases, the geometrical 'focal length' for a given aperture width was found by locating the point of least geometrical confusion. The difference between the geometrical focal length and the paraxial focus is a defocus arising from spherical aberration. To illustrate the amount of defocus that could be expected, the uniformly weighted ray-trace indicates that a positive defocus of up to 12 mm could be expected for an aperture width of 7.5 mm. The Gaussian weighted curve indicates that a defocus of 7.0 mm could be expected for the same aperture width.

The experimental results show that the amount of defocus actually observed was much less than predicted by geometrical optics. A least-squares parabola, which is a reasonably good fit is drawn through the experimental points. At an aperture of 7.5 mm the amount of defocus, read off the fitted curve, is only 2.2 mm. The evidence is that factors are present which compensate for the destructive geometrical effects. Doughty [19] postulated that a variable transmission function could reduce the amount of defocus, a conclusion borne out by the differences between the uniformly-weighted and the Gaussian-weighted ray-trace curves. The experimental data of Figure [8.31] is evidence that a tendency towards a negative focal shift, which only appears in optical propagation models based on a field analysis, counterbalances the effects of positive spherical aberration. The optical performance of a spherical geodesic lens with edge-

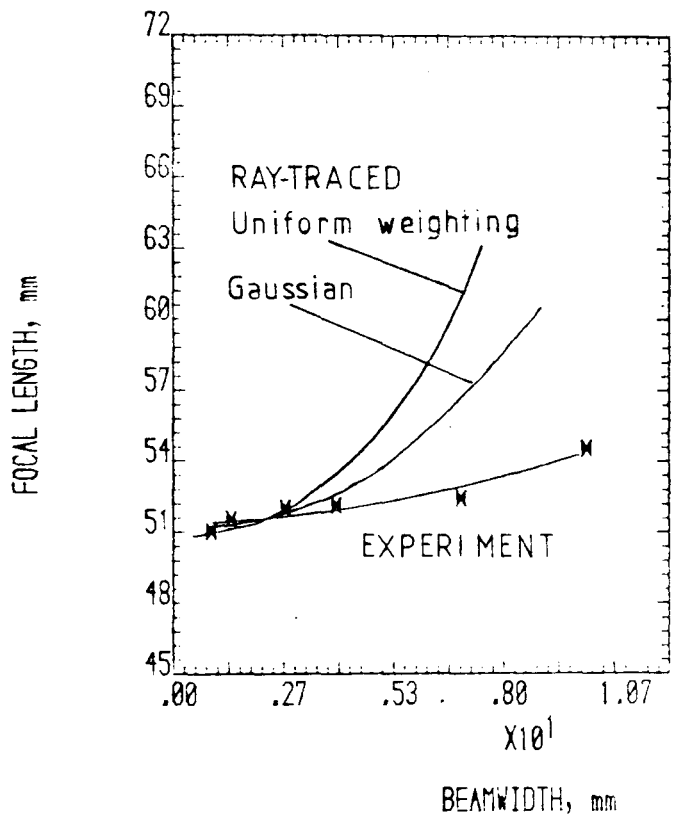


FIGURE [8,31] COMPARISON OF MEASURED AND CALCULATED VARIATION OF FOCAL LENGTH WITH APERTURE FOR LENS SL1

rounding is, therefore, better than might be expected from purely geometrical considerations.

8.12 Conclusions

Several types of inhomogeneous lens fabricated on $LiNbO_3$ substrates have been characterised in terms of observed optical performance. The lenses consisted of :

(a) a spherical geodesic lens having a toroidal edge-rounding region;

(b) an overlay lens made from arsenic trisulphide, which was a fabricated approximation to a perfect Luneburg lens;

(c) an aspheric geodesic lens fabricated as closely as possible to a theoretically perfect design;

(d) a complementary pair of aspheric geodesic lenses of perfect design, supported on a single substrate.

The spherical geodesic lens exhibited good optical properties that resulted in an ability to discriminate between two acousto-optically diffracted signals separated by 2 MHz at a centre frequeny of 746 MHz. The optical properties of this lens have been explained by postulating a balance between the positive spherical aberration introduced by the lens and the tendency towards negative defocus that is known to occur in lenses at small angular apertures.

The overlay lens exhibited poor optical performance, as had been predicted by ray-tracing techniques in chapter four. Negative spherical aberration was observed which resulted in a negative defocus of nearly 25%, in comparison with a theoretical value of 20%. The discrepancy was related to differences between the calculated approximate profile investigated in chapter four and the actual fabricated profile shown in chapter seven.

The single aspheric geodesic lens yielded diffraction-limited performance at a relative aperture of f/10 . At wider apertures the lens exhibited degraded optical performance. The width of the focal spot became several times the value of the diffraction limit, and sidelobe levels became high.

The geodesic lens pair on a single substrate exhibited different optical characteristics. One lens produced a smooth focal field, albeit wider than the diffraction-limited case. The other lens exhibited a severely degraded focal field, probably due to the presence of local scatterers in the vicinity of the focus, either in the waveguide itself or at the polished end-face of the waveguide.

REFERENCES FOR CHAPTER EIGHT

[1] Shannon,R.R.,
'The Testing of Complete Objectives', Ch. 5 of 'Applied Optics and Optical Engineering', ed. Kingslake,R., Academic Press, New York, 1965

[2] Suzuki,K, Ogura,I, and Ose,T.,
Measurement of spherical aberrations using a solid-state image sensor, Appl. Opt., 18(22), 1979, 3866-3872

[3] Vaidya,W.M. and Sen Gupta,M.K.,
Measurements of axial and off-axis geometrical aberrations of microscope objectives, J. Opt. Soc. Am., 50(5), 1960, 467-477

[4] Klein,W.R., Cook,B.D. and Mayer,W.G.,
Light diffraction by ultra-sonic gratings, Acustica, 15(2), 1965, 67-74

[5] Lean,E.G.H., White,J.M. and Wilkinson,C.D.W.,
Thin-film acousto-optic devices, Proc. IEEE, 64(5), 1976, 779-787

[6] Anderson,D.B, Davis,R.L., Boyd,J.T. and August,R.R.,
Comparison of optical-waveguide lens technologies, IEEE J. Quant. Electron., QE-13(4), 1977, 275-282

[7] Gaskell,J.D.,
'Linear Systems, Fourier Transforms, and Optics', Wiley and Sons, New York, 1978

[8] Born,M. and Wolf,E.,
'Principles of Optics', Pergamon Press, 1975

[9] Yariv,A.,
'Introduction to Optical Electronics', Holt, Rinehart and Winston, New York, 1976

[10] Yao,S.K., Anderson,D.B, August,R.R., Youmans,B.R. and Oania,C.M.,
Guided-wave optical thin-film Luneburg lens: fabrication technique and properties, Appl. Opt., 18(24), 1979, 4067-4079

[11] Marom,E., Chen,B. and Ramer,O.G.,
Spot-size of focused truncated Gaussian beams, Optical Engineering, 18(1), 1979, 79-81

[12] Goodman,J.W.,
'Introduction to Fourier Optics', McGraw-Hill, New York, 1968

[13] Cagnet,M., Francon,M. and Thrierr,J.C.,
'Atlas of Optical Phenomena', Springer-Verlag, Berlin, 1962

[14] Ross,G., Fiddy,M.A. and Nieto-Vesperinas,M.,
The inverse-scattering problem in structural determinations, in 'Inverse Scattering Problems in Optics', Baltes,H.(ed) , Topics in Current Physics, volume 20, Springer-Verlag, Berlin,1980

[15] Van Toorn,P. and Ferwerda,H.A.,
On the problem of phase-retrieval in electron microscopy from image and diffraction patterns, Optik, 47(2), 1977, 123-134

[16] Barakat,R.,
Numerically stable iterative method for the inversion of wave-front aberrations from measured point-spread function data, J. Opt. Soc. Am., 70(10), 1980, 1255-1263

[17] Maeda,J.,
Retrieval of wave-aberration from point-spread function in optical transfer function data, Appl. Opt., 20(2), 1981, 274-279

[18] Driscoll,W.G. (ed), Vaughan,W.(ed),
'Handbook of Optics', McGraw-Hill, New York,1978

[19] Doughty,G.F.,
'Aspheric Geodesic Lenses for an Integrated Optical Spectrum Analyser', Ph.D. Thesis, University of Glasgow, 1983

CHAPTER NINE

CONCLUSIONS AND RECOMMENDATIONS FOR FUTURE WORK

CHAPTER NINE

CONCLUSIONS AND RECOMMENDATIONS FOR FUTURE WORK

9.1 Review of thesis

The goals of this work were:

(i) to fabricate integrated optical Luneburg lenses of the overlay type;

(ii) to characterise, experimentally, already-fabricated integrated optical geodesic lenses;

(iii) to implement and test computer models of optical propagation through integrated optical inhomogeneous lenses, of which geodesic and Luneburg lenses are special cases. The models were (a) geometrical optical, based on ray-tracing techniques and (b) wave optical, based on the recently developed beam propagation method (BPM).

All of the above goals have been met, with varying degrees of success.

Chapter one, the introduction to the thesis, has attempted to show that integrated optical components, and lenses particularly, could play a key role in the optical processing revolution which is widely predicted. In chapter two, the historical development of inhomogeneous lenses has been reviewed. Inhomogeneous lenses have been known, for some decades, to offer the possibility of achieving near-perfect optical performance. As has often been the case in integrated optics, many useful developments were initiated at microwave frequencies.

The design principles and theory of inhomogeneous lenses have been presented. Luneburg lens refractive index profiles and

thin-film overlay thickness profiles which provide the required refractive index variation, through the dispersion of the effective propagation constant with thickness, have been computed. The lens material was arsenic trisulphide (As_2S_3), a high refractive index material supported on a lithium niobate ($LiNbO_3$) substrate on the surface of which a titanium-diffused waveguide was the ambient guiding medium. Tolerances have been established which relate the sensitivity of lens focal-lengths to overlay-film thickness variations. It is found that lenses having large f-numbers are extremely sensitive to film thickness variations. Lenses having f-numbers as low as f/1, on the other hand, display thickness profiles that are not likely to be realizable using simple fabrication techniques. A goal of fabricating an f/2 lens is a reasonable choice, given these conditions.

A theory of integrated optical geodesic lenses has been presented which is used to design lenses having an inner focusing region and an outer edge-rounding region bridging the inner region and the ambient waveguide. The composite lens is capable of perfect focusing, theoretically, within the usable inner aperture. Lenses that have been fabricated by Dr. G.F. Doughty and co-workers at the Department of Electronics and Electrical Engineering at the University of Glasgow have a full aperture f-number of f/1.85, of which only a relative aperture of f/2.5 is usable.

A simple geodesic lens has also been fabricated which has a spherical inner focusing region and a toroidal outer edge-rounding region. This lens has a full-aperture f-number of f/4.98, of which only a relative aperture of f/10 or greater is usable, due to the aberrations introduced by the spherical geometry.

The vacuum environment to be used for fabricating thin-film overlay lenses has been investigated in chapter three. Simple models of the spatial flux properties of evaporation sources, including the well-known Knudsen model, have been presented in

this chapter. The variation of the thickness of a thin film deposited on a planar substrate, held above a small Knudsen source, has been computed and compared with desired lens profiles for several source-to-substrate distances. The distances required to produce significant variation are small, of the order of a few millimetres. Furthermore, the variations do not correspond very well with the required profiles. However, the actual source used in fabrication is not small and, consequently, substantially modified evaporant flux profiles are expected. Reports in the published literature have also demonstrated that 'blocking', or 'shadow' masks interposed between source and substrate can be used to control the evaporant flux properties and, consequently, the deposited film profiles.

A computer model, incorporating a source of cylindrical geometry and a single infinitely-thin mask with a circular opening, has been developed which is used to predict the distribution of deposited material on a planar substrate. The geometry of the source is held fixed at the dimensions of the experimental source. The radius of the mask opening and the placement distances of the mask, relative to source and substrate, are used as variables in the search for the best profiles. The 'best profile' is found by minimising the least-squares deviation of the deposited profile from the required Luneburg lens. The approximate profiles obtained differ from the desired profiles in that the central regions are flat-topped, and curve steeply at the edges. Optical aberrations were expected to occur, together with high scattering losses at the edge of the lenses.

Algorithms for tracing rays through geodesic and overlay lenses have been given in chapter four. These algorithms were published by W.H. Southwell and have been implemented, with modifications, by the present author and J. van der Donk and P. Lagasse of the University of Gent, Belgium.

Comparisons between ray-traces through the conceptually perfect f/2 Luneburg lens and the best approximate profile

obtained using the method discussed in chapter three have revealed that the approximate profile does indeed give rise to spherical aberration. The effective focal length is reduced by 20% with respect to the design focal length, at half-full aperture. Rays outside the half-aperture are deflected at very large angles with respect to the axis of optical propagation, as a result of the steepness of the profile edges.

Comparisons between ray-traces through the theoretical aspheric geodesic lens and two achievable approximations have revealed that small spherical aberrations occur which oscillate rapidly as a function of the ray position in the aperture. The aberrations, being small, are unlikely to degrade the optical field significantly. Furthermore, the perturbations of the lens profile which give rise to the aberrations are expected to be greatly reduced by polishing.

Ray-traces through the spherical geodesic lens with a rounded edge have indicated that the lens performance is significantly degraded by spherical aberration at wide apertures. Near full-aperture, a positive defocus of nearly 40% is required to locate the point of least confusion.

The geometrical optics techniques used in chapter four are insufficient to describe fully optical propagation through inhomogeneous lenses, since diffraction effects can not be modelled. In chapter five, a numerical method for propagating optical wave-fields through inhomogeneous lenses has been introduced. The method is known as the beam-propagation method (BPM) and relies heavily on the techniques of Fourier optics. The theory and practical implementation of the technique has been discussed in chapter five. The technique had been implemented by J. van der Donk and P. Lagasse, but had not been extensively tested. The BPM software was made available to, and modified by, the present author.

Obtaining correct results under diffraction-limited conditions represented a rigorous test of the capabilities of

the BPM. The aspheric geodesic lens of theoretically perfect design presented in chapter two has been used to evaluate these capabilities. The software has been modified to give detailed information about the amplitude of the field in the region of the focus. The method has yielded generally excellent results, for a variety of incident field conditions. The amplitude and energy distributions closely resemble the classical results published in the literature and the striking differences between the focal fields arising from uniformly illuminated apertures and truncated-Gaussian illumination have been highlighted.

A negative focal-shift has been observed with the BPM that has also been predicted by other recently-published diffraction analyses. The focal-shift obtained was comparable in magnitude to the published values. The focal-length of a spherical geodesic lens having a rounded-edge was shown to be less than that predicted by raytracing at a limited aperture. It is speculated that the tendency of the focus to move towards the lens will compensate, to an extent, for the positive geometrical aberrations of the spherical geodesic lens having a rounded edge.

In chapter six, the fabrication of planar waveguides and overlay lenses has been discussed. The properties of lithium niobate, titanium-diffused waveguides and arsenic trisulphide have been reviewed. Arsenic trisulphide films are known to benefit from irreversible annealing, in terms of optical and mechanical stability. A reversible annealing process has been reported which could allow for post-fabrication modification of arsenic trisulphide waveguide properties.

Fabricated titanium-diffused waveguides have performed well, with low in-plane scattering. In contrast, planar waveguides of arsenic trisulphide have performed quite poorly, exhibiting high levels of loss and in-plane scatter.

It has been found that accurate control of the evaporation process used for fabricating arsenic trisulphide overlay lenses is impossible, due to the very small distances required for

achieving good profile approximations. Nevertheless, a measured lens profile has been presented which comes close to the desired profile. An alternative technique for the fabrication of overlay lenses has been proposed.

In chapter seven, the problems of achieving good uniformity of titanium-diffused waveguides over geodesic lens surfaces have been discussed. It has been shown that a simple Knudsen source will yield a very uneven film of deposited titanium, which will influence the diffused waveguide in similar fashion. The degree of non-uniformity is especially severe for the highly-curved aspherical geodesic lenses. It is postulated that such non-uniformity will affect the performance of these lenses adversely, influencing both the amplitude transmission properties and the phase-transformation properties.

The optical properties of fabricated lenses have been reported in chapter eight. A single thin-film overlay lens approximation to a Luneburg lens, a spherical geodesic lens with a rounded edge, a single aspheric geodesic lens, and an aspheric geodesic lens pair have been reported upon. The thin-film lens performs as predicted by ray-tracing, ie a negative defocus arising from spherical aberration occurrs. The spherical geodesic lens performs surprisingly well, and is able to discriminate successfully between two closely-spaced acousto-optically-diffracted signals. The focal-length of the lens varies much less as a function of aperture than is predicted by ray-tracing, in confirmation of a postulate arising from the BPM wave-optics analysis.

The single aspheric geodesic lens yields diffraction-limited performance at a relative aperture of f/10. At lower relative apertures the focal spot-size remains approximately constant. The sidelobe levels, however, are generally rather high and asymmetrical, possibly indicating an asymmetrical deposition of titanium before waveguide fabrication. The complementary pair of aspheric geodesic lenses exhibits different levels of optical performance, but both lenses perform significantly worse than the

single lens, which may be a further indication that waveguide uniformity is an important problem.

9.2 Conclusions on overlay Luneburg lenses

The overlay thin-film lenses perform poorly for the following reasons:

(a) the choice of As_2S_3 as the overlay material, dictated by its high refractive index with respect to titanium-diffused $LiNbO_3$, necessarily lowers expectation in other respects, notably in terms of in-plane scattering loss, absorption loss, mechanical and optical stability,

(b) the need to improve the stability of the films by annealing, therefore raising the refractive index from the as-evaporated state to a value close to the bulk state, shifts the curves of focal-length sensitivity to film thickness changes to regions of very low f-number. Unfortunately, the regions of low f-number are just the regions where the overall lens profiles are most difficult to manufacture using simple techniques;

(c) the profiles require very short source-to-substrate and source-to-mask distances, as calculated by the shadow-masking model. Such distances preclude good control or on-line measurment of fabricated films;

(d) the steep edges of the approximate profiles lead to extremely high values of scatter at the ambient waveguide/lens interface, leading to very large throughput losses.

9.3 Recommendations for future work on overlay lenses

Several recommendations can be made that would lead to much improved overlay Luneburg lens performance:

(i) the optical problems associated with As_2S_3 would be

alleviated by operating at a longer wave-length, in the near infra-red. The losses associated with scattering and absorption would decrease. The optical characteristics of the lenses at infra-red wavelengths, however, being invisible to the eye are more difficult to assess.

(ii) Following the success of Busch and co-workers with unannealed films (see chapter two for reference), annealing should be eliminated thereby reducing the refractive index difference between the lens and the surrounding waveguide. Such a reduction decreases the sensitivity of the focal length to changes in film thickness at larger f-numbers where the profiles will be flatter and, consequently, much easier to fabricate. However, a discontinuation of the annealing process raises questions about the long-term stability of the lenses. Other materials having a slightly lower refractive index than As_2S_3 could also be investigated.

(iii) Alternatively, if annealed films are necessary, the effective refractive index of the surrounding waveguide should be raised. This could be done by using As_2S_3 itself as the ambient waveguide material. However, the other problems already mentioned would work against such a solution. A more exciting approach would be to use the constantly improving waveguide technology of proton-exchange which allows well-controlled refractive index changes of up to 0.12 at visible wavelengths.

(iv) Profiles based on the dispersion of the effective propagation constant of TM modes, rather than TE modes, tend to be flatter and are therefore relatively easier to fabricate as has been shown by Busch and co-workers.

(v) Sputtering techniques, which generally involve much larger source areas than evaporation techniques, would allow shadow-masking models with a larger 'choice' of regions to selectively block off from the substrate. A much-improved profile fit would result.

(vi) The dimensions of the cylindrical source in the shadow-masking model used here were held constant during the computer search for a close-fit profile. It would be useful to allow these dimensions to vary to see whether closer fits could be obtained.

(vii) A more sophisticated shape of mask than that used in the present work has been reported by Yao and co-workers, and also by Hatakoshi and co-workers (see chapter two). These masks repay the costs of development and fabrication because, once the mask is fabricated, the fabrication of lenses becomes both cheap and reproducible.

(viii) The computer-controlled, motor-driven iris method proposed in chapter six would be extremely interesting once set-up, since changes to the technique would be programmable.

(ix) Finally, a theoretical investigation of Luneburg lens profiles that did not possess abrupt transitions between the surrounding waveguide and the lens would repay investigation. Bridging regions, analogous to the rounded edges of the geodesic lenses investigated in this study, would reduce scattering and reflection loss at the interface.

9.4 Conclusions on geodesic lenses

Spherical geodesic lenses with toroidally rounded-edges would appear to be reasonable candidates for low-to-medium performance specifications, in view of the optical characteristics measured in this work. The small change in focal length as a function of aperture, in comparison with the shift predicted by geometrical optics, is all the more surprising in view of the fact that the spherical aberration data measured using a Hartmann experimental ray-trace corresponded closely to the theoretical curves. This may be taken as confirmation of the fact that the wave-optical negative focal shift does not depend on the geometrical characteristics but rather on whole-aperture field interference effects.

The large, unevenly distributed amounts of optical energy scattered into the regions outside the central spot in the aspheric geodesic lenses are indications, once again, that the problem of in-plane scattering is still one of the most pressing in integrated optics. Waveguide uniformity has shown itself to be a further problem, as demonstrated by the failure of the central regions of the complementary lens pair to transmit light after an initial waveguide formation. For these reasons, the aspheric geodesic lenses are falling short of diffraction-limited performance at wide apertures.

9.5 Recommendations for future work on geodesic lenses

(i) Given the improvements in proton-exchange waveguides, especially those fabricated in dilute melts, it is probably time to try this waveguide technology as a competitor against titanium-diffusion.

(ii) The variation of deposited thickness of titanium across geodesic lenses should be measured to see how the experimental data compares with the theory given in chapter eight. The transmission properties of the lenses as a function of aperture should be measured and a study carried out on whether a correlation exists between the waveguide uniformity and the transmission function.

(iii) Lenses should be investigated under conditions of off-axis incidence. The problem of anisotropy, which has been largely ignored in the present study, would almost certainly become more pressing under these conditions. One of the most easily controllable and accurate methods of varying the incidence angle is also one of the most directly relevant, namely the Bragg interaction of surface-acoustic waves with an optical beam.

(iv) The possibilities of modifying geodesic lens properties should be investigated with regard to decreasing in-plane scatter, apodizing to reduce sidelobe levels and changing the

focal length. The latter property is especially important if the edges of waveguides are to be close to the foci. The BPM technique shows that the depth of field at wide apertures is very small, so that the field can decay rapidly at very short distances from the achieved focus. Post-fabrication modification of the focal length would enable the beam to be tuned to correct spatial positions. One way of doing this would be to introduce loading layers of overlay material.

(v) Experimental observations of negative focal-shifts in diffraction-limited circumstances, such as those predicted by the BPM and other diffraction analyses, are not known to the author. It would be extremely interesting to obtain such results.

9.6 Conclusions on optical propagation models

The geometrical optics technique of ray-tracing was found to give useful information about large aberrations, such as those exhibited by the overlay approximation to the Luneburg lens, and the spherical geodesic lens with a rounded edge. It was found more difficult to interpret the effects of vary small aberrations such as those exhibited by the achievable approximation to the aspheric geodesic lens. The main utility of ray-tracing is to obtain a relatively crude idea of the optical performance of a given lens system. Where aberrations are large, such as in low-performance applications, ray-tracing is sufficient to describe lens performance.

The BPM, on the other hand, appears to be a very powerful technique for ascertaining optical performance in the form which is directly measured. Some uncertainty still exists about the accuracy of the focal shift predicted by the method, although there is no doubt that negative focal shifts are a feature of systems that are perfect from a geometrical optics point of view.

9.7 Recommendations for future work

(i) Ray-tracing should be used in conjunction with shadow-masking programs to investigate other overlay lens profiles that may provide for better approximations to the desired profiles. The profiles would then be optimised optically, rather than in terms of physical closeness to a desired shape.

(ii) Further tests should be carried out to establish the accuracy of the beam propagation method. These should consist of numerical tests for convergence and more detailed comparisons with other theories and empirical data.

(iii) Phase information is available with the BPM, since the complex amplitude of the wave-field is calculated. The phase behaviour of the field in diffraction-limited systems should be investigated and compared with that obtained using classical theories. In particular, the relationship of the phase-fronts to those predicted by geometrical optics should be examined in light of the focal shift.

(iv) The BPM should be used to investigate off-axial incidence conditions, the effects of aberrations on the optical field in the focal region, the effects of profile perturbations and apodisation effects. An anisotropic version of the BPM is available which should be tested for validity. Furthermore, the BPM can be used to model acousto-optic diffraction but has been under-utilised in this respect. It is particularly important to investigate inter-modulation effects, which can degrade the resolution of optical processors.

9.8 Summary of original work contained in this thesis

This thesis has reported original work in at least eight respects:

1. The thickness profiles and fabrication tolerances of overlay Luneburg lenses composed of anneal-stabilized films of As_2S_3, deposited onto Ti:$LiNbO_3$ waveguides, have been calculated.

2. The actual form of the refractive index profiles of the Ti-diffused waveguides has been shown to be unimportant for the shape of the overlay Luneburg lenses when a high-index film is used to fabricate the lenses.

3. An original model and computer optimization method has been developed for the problem of fabricating Luneburg lenses in vacuum evaporation and shadow-mask conditions.

4. The mathematics of a well-known algorithm for tracing rays in Luneburg lenses has been corrected and properly established.

5. A significant focal shift from the geometrical optics value has been calculated using a wave propagation method. This is the first time that such a shift has been predicted in an integrated optics context.

6. Waveguide uniformity problems on geodesic lenses have been investigated from a theoretical point of view, for the first time in integrated optics.

7. A range of optical assessment methods have been borrowed from bulk optics and used extensively to good effect in testing integrated optical lenses.

8. The spherical geodesic lens with a rounded edge has been shown to be more competititive with other types of integrated optical lens than had been previously thought.

Note (1) re. Equation (4.35), Chapter Four

The difference between the expression obtained by the present author for the second derivative of the Luneburg lens refractive index profile, and that obtained by Southwell (shown below) is quite important. In spite of the errors in his expression, Southwell presented ray-tracing results which, ostensibly, were based on the published algorithm. Since, however, the expression gives ridiculously high values of the second derivative which would normally cause the algorithm to fail, it must be assumed that Southwell himself did not utilise the expression as published and relied on numerical calculations of the derivative instead. It would be expected that the expression given in (4.35) would lead to more accurate results than those obtained using numerical techniques. This is indeed found, and a gain in computing efficiency is also obtained.

Southwell also got his differentiation mixed up in a paper on geodesic lenses ("Geodesic optical waveguide lens analysis", J.Opt.Soc.Am., 67(10), pp 1293-1299) as was pointed out subsequently by Marom and Ramer (E. Marom and O.G. Ramer, "Geodesic optical waveguide lens analysis: Comment", J.Opt.Soc.Am., 69(5), pp 791-792) and acknowledged by Southwell (same issue, pp 792-793). The present author is not aware, however, of a published correction to the mistake reported in the present work.

Equation (14) from Reference [11] of Chapter Four:

$$\frac{d^2n}{dr^2} = \frac{2n^3\left(G^2(1-rn)^{1/2} + H\right)}{\left[2(1-rn)^{1/2} + nrG\right]^3} + \frac{2}{n}\left(\frac{dn}{dr}\right)^2$$

PUBLICATIONS

[1] " Integrated optical microwave spectrum analyser (IOSA) using geodesic lenses. "
G.F. Doughty, R.M. DeLaRue, N. Finlayson, J. Singh and J.F. Smith
Proceedings of the Society of Photo-Instrumentation Engineers, volume 369, 1982, 705-710

[2] " Integrated Optical Techniques for Acousto-Optic Receivers. "
S.M. Al-Shukri, A. Dawar, R.M. DeLaRue, G.F. Doughty, N. Finlayson and J. Singh
Chapter eight, ' Advanced Signal Processing', edited by D.J. Creasey. Peter Pergrinus, Stevenage, Herts, UK, 1984

[3] " Acousto-optic techniques in integrated optics. "
A. Dawar, R.M. DeLaRue, G.F. Doughty, N. Finlayson, S.M. Al-Shukri and J. Singh
Proceedings of the Society of Photo-Instrumentation Engineers, volume 517, 1984, 64-73

[4] " Integrated Optic Lenses. "
N. Finlayson and R.M. DeLaRue
Post-deadline paper presented at 1985 Summer School on Surface Excitations, Erice, Sicily. To appear in proceedings of School.